Essays in Social and Behavioral Economics

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Brian Louis Zurowski

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Jan Werner, Advisor

September, 2015

Acknowledgements

I wish to thank my advisor, Jan Werner for his constant support and faith in my ability to succeed. From our conversations, the wisdom he imparted, and his attention to detail, I have sharpened my focus and philosophy. I am grateful to Beth Allen, who consistently provided feedback on ideas I presented in workshop and fostered a habit of asking how things might look to others. I thank David Rahman for his positive energy, humility, and appreciation of different writing styles. I also thank Kim Sau Chung, Aldo Rustichini, and Terry Hurley for their guidance, especially in the early years of my graduate studies.

I wish to thank Simran Sahi for giving me the opportunity to push myself as a teacher in many interesting roles. I am grateful for the assistance and advice of Caty Bach and Wendy Williamson. I also thank Ellen McGrattan, whose calm rationality and presence during the job market gave me the strength needed to persevere.

I thank my friends and colleagues at the University of Minnesota, notably Andrew Dust, Ruchika Mohanty, Rene Schwengber, Murat Ali Cengelci, and all the participants of the Math Econ workshops, as well as Minesh Amin, Ramses Armendariz, Michael Maio, and Eric McDermott. I thank Pei Cheng Yu for sharing his knowledge of the literature on temptation, and Kelvin Wong for contributing his talents and time to investigate an empirical question related to social dissonance.

Many, many thanks to my family and friends at home and elsewhere for their patience, understanding, and love. I would be nothing without them.

Abstract

This thesis studies the effects of the conformity motive and temptation on individual decision-making. Social dissonance is the discomfort of choosing an action different from what others have chosen. A game-theoretical framework is used to study situations where social dissonance influences behavior or expression of opinions. Each individual may have different intrinsic preferences but is affected by the same "institution" which is modeled by a social dissonance function that evaluates the negative effect of disagreement with others. Equilibria of social dissonance games have properties such as monotonicity of choices with respect to intrinsic preferences, and monotone comparative statics with respect to changes in intrinsic preferences and institutions. "Impulse and Temptation" acknowledges that consumers who purchase larger packages of certain goods are tempted to consume more than originally needed. The analysis attempts to understand policies that prohibit the purchase of smaller packages, exploring issues of consumer naivete about temptation and addiction. It makes little sense to restrict access to small packages in a one-period model with or without consumer naivete, but it is possible that in a multi-period setting, defense against addiction is a valid reason for prohibiting sale of small packages of tempting goods.

Contents

List of Figures

Chapter 1

Introduction

Our conscious decisions are influenced by hidden factors, both internal and external, that all the accumulated science of the world is still struggling to understand. Developments in the past few decades have been questioning and defining the boundaries between conscious and unconscious, individual and social. As consumers and citizens, we behave in ways that are connected to the people and objects around us. The essays in this thesis examine interactions between individual preferences and outside influences, and how policy can either intensify or diminish the power of those influences.

Social dissonance is the discomfort of choosing an action or opinion different from others. The level of discomfort from disagreement is in part determined by the attitude set by institutions: messages from laws, leaders, and language correspond to a social dissonance function. Just as the people in one's society can affect decisions, so can the goods that we purchase and possess. In "Impulse and Temptation," I ask whether it is a helpful policy to prohibit consumers from purchasing small quantities of goods in order to satisfy an impulse, when purchases of larger quantities feed the temptation to consume further.

The first two essays discuss games with social dissonance. Chapter 2 introduces the model of Social Dissonance Games in a finite-agent setting, focusing on the study of two-action games. I present a model of individual choice within a society where conformity enters the decision-making process. Conformity for its own sake can be a powerful motive, even when one's intrinsic character is crying out to express itself in a different way. In certain situations, intrinsic preference is in conflict with social norms, and the individual experiences social dissonance when choosing differently than others. I characterize the equilibrium sets of social dissonance games and describe conditions under which *true-to-type* behavior is an equilibrium. In general, I show that equilibrium action profiles obey monotonicity: arranging agents in descending order of their liking of one action relative to its alternative, agents up to some threshold will choose the first action and the rest of the agents will choose its alternative. I then establish the supermodularity, or strategic complementarity, of social dissonance games, which allows us to use the results of Milgrom and Roberts (1990) to guarantee existence of a pure strategy equilibrium and monotone comparative statics properties. The equilibrium sets are monotone with respect to players' intrinsic utilities in that when intrinsic utility for one action increases relative to the other, equilibria shift toward that favored action. Likewise, when the social dissonance function for one action decreases relative to the other, equilibria shift toward the action with lowered dissonance.

Individuals may weigh others' attitudes on an aggregate level, rather than on a person by person basis. In a large population, this means that one ordinary individual cannot influence behavior single-handedly. When a widespread change in opinion would otherwise occur, a switching cost may prevent the coordinated shift. To handle these possibilities, Chapter 3 provides extensions of the Social Dissonance model to allow for a continuum of agents and dynamics, and discusses use of the model in empirical analysis. The main results from the finite-agent model - monotonicity with respect to intrinsic utilities, supermodularity, and monotone comparative statics - carry over to the continuum model. One additional artifact of the continuum model is that asymmetric equilibria exist where agents with a clear intrinsic preference for one action over the other can be divided into two groups that choose different actions. The continuum of agents model lends itself well to empirical applications. Assumptions about the continuity of social dissonance functions and the distribution of player types may allow for a closed-form solution to such games, especially when the specification is linear.

In Chapter 4, "Impulse and Temptation," I address a policy question about the packaging of unhealthy products such as cigarettes, and focus on the internal forces of temptation and addiction. Specifically, I ask whether the removal of a small package from the choice set inhibits or promotes excess consumption. Gul and Pesendorfer (2001) among others have noted that the mere presence of tempting alternatives alongside one's optimal choice can actually decrease the utility of an agent. A tempted agent must exercise costly self-control or else succumb to a harmful choice. I present a model of Impulse and Temptation where agents facing an impulse to consume choose from a set of packages with different quantities of the same good. After the initial impulse, choice of package, and initial consumption, agents may face an additional impulse to consume. The probability of this follow-on "temptation" impulse is increasing in the size of the package purchased, so agents may wish to purchase a small package in order to reduce temptation. It is clear that in the one-period model with sophisticated consumers, removing the small package from the choice set can only reduce welfare. Further, consumer naivete about the strength of temptations does not seem to justify prohibition of small packages. In a two-period model, it is possible that removal of small packages may be justified by a process of addiction where consumption leads to greater impulses in the future.

Chapter 2

Social Dissonance and Individual Choice

2.1 Introduction

Individuals, and the decisions they make, are rarely isolated. As consumers and business owners, we realize that our actions will be observed and judged by others, and we learn to judge our own actions as we think others in our society would. We have different preferences but are all subject to the pressure to do as others do. In a society full of many others who choose an action in line with our wishes, a decision is easy. However, a battle takes place in one's mind when society's choices clash with one's own preferred action. *Social dissonance* is the discomfort or disutility of choosing an action different from others.

In social dissonance games, agents have intrinsic preferences regarding their own actions and also care about what others are doing. The intrinsic preferences can capture the pecuniary value of different actions, the emotional attachment to different alternatives, or deeply-held moral values. Agents' payoffs also depend on how many others choose the same action. When others choose differently, a cost, *social dissonance* is incurred. A social dissonance function captures the impact of institutions, including laws, leaders, language, and culture.

The effect of social dissonance manifests itself in opinion polls, market behavior, and social norms. When our political opinions are elicited, even in the absence of company, we consider our internal attitudes and the likely opinions of others in society. In a survey, support for a particular issue is based on what we think and also what we think we should think. These *shoulds* are informed by our beliefs about others' attitudes - would they agree or disagree? If our opinions were made public, how would that impact others' opinions of us?

Public support for same-sex marriage in the United States reached the point where a majority of Americans favor its legalization, with support increasing from 40% in 2009 to 55% in 2014.[1] (Gallup, 2014). While a small percentage of Americans have a personal stake in the legal status of same-sex marriage[2] , many know someone who would benefit from legalization. Another compelling reason to favor same-sex marriage is that many others in the community also favor it. Likewise, if others in the community oppose it then we may feel compelled to also express opposition.

Since the year 2010, American institutions have shifted in their treatment of same-sex marriage. Most notably, the law has made progress toward giving same-sex couples the same privileges as heterosexual couples. Leaders have spoken in support of the rights of same-sex couples. Additionally, there has been a shift in the language of the conversation. All of these factors have reduced social dissonance for supporters of same-sex marriage, making it easier to express their opinions, even when in the presence of implied dissent.

Social dissonance is weighed in financial decisions as well. One such decision is whether or not a homeowner should default on a housing loan. In some jurisdictions, borrowers are allowed to walk away from their mortgages without recourse for the lender. Even when recourse is available, high legal costs can discourage lenders from suing defaulted borrowers.

[1] See figure 1 in appendix B for the data series.

[2] The percentage of Americans who identify as gay, lesbian, bisexual, or transgender (GLBT) is estimated at around 4%.

Of course, the borrower's future ability to obtain credit will suffer, but some may feel it is worth that cost to escape the debt. *Strategic defaults* are those in which the borrower has adequate financial means, but discontinues payments because the value of the mortgage exceeds the value of the house.

Guiso, Sapienza and Zingales (2009) consider three groups of factors important in the decision-making process of the borrower. Firstly, the borrower compares the pecuniary value of defaulting to that of continuing to pay the mortgage; if the market value of the house is lower than the remaining amount to be repaid, the borrower may wish to default. The damage to the borrower's credit rating and subsequent reduction in access to credit is also contained in the category of pecuniary considerations, which, in the model to follow, would be classified as intrinsic utility. Secondly, the borrower must weigh the relocation costs that would follow a default. Lastly, there are "moral and social considerations." Moral considerations can be defined as those deeply-held beliefs which are not based on others' observation of one's behavior, and their value should be included in intrinsic utility. As one would expect, subjects in Guiso et al.'s survey who said it is immoral to strategically default were less likely to say they would do so themselves. Social considerations are the costs resulting from others' observations and actions. When choosing an action different from others, we may feel guilt or shame; these emotions are more strongly felt when a greater number of people take a different action. Conversely, negative emotions associated with an action are reduced when more people take that action. Guiso et al. found that survey respondents who knew someone who strategically defaulted on a mortgage are more likely to express the willingness to do so themselves.

The consumption of certain goods can be influenced by social dissonance. U.S. institutions have aggressively pushed an anti-smoking agenda since the mid-1990's. This has been motivated in part by the health hazard of secondhand smoke, of which the public's estimation remained roughly constant from 1997 to 2007. However, over that same time period, public support for complete bans on smoking in restaurants rose from 40% to 54% (Gallup, 2007). Smokers represent roughly one-fifth of the adult population. In another survey by Gallup, 47% of smokers said in 2007 that they feel unjustly discriminated against, up from 32% in 2001. Restricting and taxing smoking may have the benefit of preventing citizens from becoming smokers and encouraging others to quit, but for some existing smokers, restrictions do little but expose them to feelings of shame and marginalization. Citizens who would otherwise choose an unpopular action can be "entrapped" by the high social dissonance sanctioned by institutions. Worse yet, with intense enough social dissonance, the unique equilibrium may be one in which some conceal their true opinions and everyone acts the same way. People may sacrifice voicing their own opinions in order to feel included (Noelle-Neumann, 1974).

Social dissonance games provide a framework for weighing the relative importance of intrinsic preferences and the conformity motive in decision-making, and can be used to predict the response of public actions and opinions to a change in institutions. Using this framework, I characterize the equilibrium set, and describe the separate effects of shifts in intrinsic preferences and social dissonance functions. I show that two-action social dissonance games belong to a larger class, supermodular games, in which strategic complementarities exist among agents' choices. As more agents coordinate on a particular action, the reduced social dissonance associated with the action makes it more attractive to an agent that might choose otherwise. I apply the results of Milgrom and Roberts (1990), which guarantee the existence of a pure-strategy equilibrium and provide some monotone comparative statics properties. Specifically, if for all agents there is a non-decreasing change in the intrinsic utility of one action relative to the other, the equilibrium set shifts toward profiles where more agents select the former action. Conversely, if dissonance for one action increases relative to the dissonance for the other action, the equilibrium set shifts toward the latter action.

There are some games where intrinsic utility for a preferred action is strong enough relative to social dissonance that choosing one's preferred action is a dominant strategy.

If this is the case for all agents, there is a unique *true-to-type* equilibrium. If only some agents have such strong intrinsic preferences, the other agents may decide to follow them in choosing the same action; I call this a *monolithic* equilibrium. If more than one equilibrium action profile exists, agents may fail to coordinate on the most efficient one.

In general, I show that equilibrium action profiles obey monotonicity: arranging agents in descending order of their liking of one action relative to its alternative, agents up to some threshold will choose the first action and the rest of the agents will choose its alternative. The exception to this is that the group of agents who are intrinsically indifferent between two alternatives may make different choices. If the indifferent agents split into two groups choosing either action, this reduces social dissonance for everyone else. However, if they all choose alike, this only reduces social dissonance for one group.

In a social dissonance game, if institutions remain fixed and beliefs about the opinions of others do not change, a shift in the equilibrium action profile can be attributed to a change in agents' intrinsic utilities. If, however, we observe a change in opinions without a change in information, there is reason to suspect that a shift in the social dissonance function was a key factor. When there is a monotone shift in society's intrinsic utilities to favor one action, the equilibrium will shift toward an action profile where more agents choose that action. Similarly, when the dissonance function for one action increases relative to the dissonance of the other action, the equilibrium set shifts away from the action with the increased dissonance.

When answering opinion polls, social dissonance is often a salient concern, especially when respondents are constantly asked for their opinions, and offered no new relevant information aside from signals about the opinions of others. If there is a decrease in the social dissonance associated with an opinion, through a public awareness campaign or a statement of support by a public leader, we would expect the expressed opinions of the population to shift in the same direction.

In a decision to default on a mortgage, pecuniary concerns weigh heavily, and unless there is a campaign to shame defaulters, maximum social dissonance is small relative to the pecuniary concerns of distressed homeowners. Still, if only a small percentage of the population defaults, the social dissonance of defaulting is nontrivial and borrowers would not casually do so because of some small perceived financial advantage. Only those with the most dire financial circumstances would choose to default. If for some reason the stigma associated with the behavior is reduced by its increased incidence, an institution can counter the trend by policies and official statements that increase the social dissonance of strategic defaults. If, on the other hand, a small segment of the population is feeling entrapped by public opinion and this is reducing overall welfare, institutions can adjust to lower social dissonance or to reduce efforts to convert otherwise indifferent agents to the majority opinion.

The paper proceeds as follows: Section 2.2 discusses related literature. Section 2.3 defines social dissonance games and dissonance functions. Section 2.4 discusses supermodularity in two-action games and provides a characterization of equilibria, while section 2.5 summarizes comparative statics results. Section 2.6 provides some examples, section 2.7 discusses cascades and entrapment, and section 2.8 addresses games with more than two actions. Section 2.9 concludes.

2.2 Literature

The experimental paradigm of Solomon Asch (1951) demonstrates the effect of conformity pressure on an individual's choices. Subjects were asked to complete a simple task of matching a reference line with one of three other lines. The error rate of isolated subjects was less than 1%, but when asked to respond after a group of confederates unanimously gave an incorrect answer, subjects made an error in more than one out of three trials.

In economics, the tradition of binary-choice games with externalities dates back to at least 1973 in Thomas Schelling's model of aggregate behavior in a binary-choice setting with externalities. Extending this in *Micromotives and Macrobehavior* (1978), Schelling describes many situations where the payoffs to agents depend on the choices of others, including a multiperson prisoner's dilemma and a coordination game. Dybvig and Spatt (1983) provide a model where agents have heterogeneous intrinsic preferences and, like Schelling (1973) have their payoffs affected by how many people adopt a certain norm.

The games discussed in this paper are similar to those specified in Dixit (2003), though the framework here is more general. Dixit specifies linear variation in intrinsic utility and externalities, and argues that "more general functional forms would generate conditions that are essentially tautologous." However, there is something to be learned by including parameters describing the changing marginal effect of concordant agents and the elimination of dissonance. Knowing these parameters can help make predictions about, or at least put bounds on, future behavior.

The positive externality from concordant actions, which can also be viewed as a negative externality from discordant actions, has been referred to by many names in social science literature, each with its own connotations. Others have examined the effects of stigma (Athreya 2004), conformity (Bernheim 1994), network effects (Farrell and Klemperer 2007; Easley and Kleinberg 2010), herding (Banerjee 1992), informational cascades (Bikhchandani, Hirshleifer, Welch 1992), and social customs (Akerlof 1980).

There is a large literature, created mostly in the early 1980's, dealing with the adoption of technological standards and the *network effects* that incentivize a society to choose only one of two or more competing technologies. For example, Farrell and Saloner (1985) address the coordination problem of switching to a new technology. A social dissonance game is a reframing of the coordination problem, with a negative externality that shrinks to zero as more players choose the same action.

Some of the spirit and language of Bernheim (1994) closely matches my own. Bernheim's model explains the conformity motive by adding an intermediate measure of *esteem* or approval of others. Esteem depends on others' *perceptions* of one's preferences, which in turn are deduced from observing actions. When esteem is sufficiently important relative to intrinsic preferences, individuals conform to a single, homogeneous standard of behavior. Even so, agents with extreme preferences refuse to conform.

Shaw et al. (1999) describe the process by which individuals use media to learn the agenda of a community, in order to "remove the social dissonance of being alone and/or dealing with ambiguity." The term *social dissonance* is chosen to indicate that the individual is negatively affected by choosing an action or opinion that is dissonant with the actions and opinions of others. In order to apply the model to a real-life situation, we do not need to understand the particular motive behind the preference for choosing a concordant action. We must only be convinced that social dissonance exists.

2.3 Social Dissonance Games

Suppose there are N agents in a society and each agent $i \in \{1, 2, ..., N\}$ chooses an action s^i from finite strategy space S^i. Action profile $s = (s^1, s^2, ..., s^N)$ lists the choices of the agents and s^{-i} lists the choices of all agents except i. Agents' payoffs are captured by utility functions $\{U^1, U^2, ..., U^N\}$ which are separable into two components: intrinsic utility and social dissonance. The intrinsic utility, $u^i(s^i)$, depends only on the agent's own action, s^i, and can be particular to each agent. Complete information is assumed.

The cost due to social dissonance, d, depends on the agent's action and the number of other agents choosing that same action. It is a mapping[3] $d : S^i \times \mathbb{R}_+ \to \mathbb{R}$. A single agent can influence the social dissonance of other agents by his choice of action but the agent influences his own social dissonance by choosing which group to join.

Agents have identical choice sets; otherwise the idea of social dissonance may not apply to agents who are unaware of or unable to choose actions that are available to other agents. For action profile s, let $N(s^{-i}, s^i)$ be the number of agents other than i who play action s^i. The dissonance function can be written as $d(s^i, N(s^{-i}, s^i))$.

[3] In finite-agent models, the second argument will be an integer in $\{0, 1, ..., N-1\}$, but defining $d(s^i, \cdot)$ over the real numbers allows for the ability to generalize.

Assembling the intrinsic utility and social dissonance, the payoff for agent i of action profile s is

$$U^i(s) = u^i(s^i) - d(s^i, N(s^{-i}, s^i)) \tag{2.1}$$

Each agent wishes to choose an action that is a best response, maximizing his payoff, given the choices of others. $s^i \in S^i$ is a **best response** to the actions of the other agents, s^{-i}, if for $t^i \in S^i$ with $t^i \neq s^i$, $u^i(s^i) - d(s^i, N(s^{-i}, s^i)) \geq u^i(t^i) - d(t^i, N(s^{-i}, t^i))$. This best response condition may be rewritten as

$$u^i(s^i) - u^i(t^i) \geq d(s^i, N(s^{-i}, s^i)) - d(t^i, N(s^{-i}, t^i)) \tag{2.2}$$

The inequality framed in this way states that the incremental benefit in intrinsic utility from choosing s^i must not be outweighed by s^i's incremental social dissonance cost.[4]

A **Nash Equilibrium** is an action profile, s, where each agent's action is a best response to the actions of the other agents. In this paper, attention is restricted to pure-strategy Nash Equilibria.[5] It follows from observation of (2.2) that the best response and the equilibrium set depend only on the *differences* in each agent's intrinsic utility for different actions and the differences in social dissonance for each profile of actions.

2.3.1 Dissonance Functions

Studies on conformity, including Asch (1951) and Gerard et al.(1968), imply that social dissonance decreases as the number of agents choosing the same action increases, and these decreases in social dissonance are greatest with the addition of the first few agents who choose the same action. When the disagreeing group becomes small enough, dissonance reaches zero. It is safe to assume there is no social dissonance associated with choosing an

[4] If both are negative, the loss in intrinsic benefit must not be outweighed by the reduction in social dissonance cost.

[5] This restriction still allows for agents with the same intrinsic utility to choose differently.

action that everyone else chooses. These two assumptions will be imposed throughout the paper:

Assumption 1. *For each player i and each action s^i, $d(s^i, \cdot)$ is non-negative and is non-increasing in the second argument.*

Assumption 2. *When all other agents choose an action that is concordant with the choice of a given agent i, i will experience no social dissonance. That is, $d(s^i, N-1) = 0$* [6] *for each player i and each $s^i \in S^i$.*

One additional assumption on the dissonance function which may be imposed, and can make analysis simpler is the following:

Assumption 3. *The social dissonance function, $d(s^i, \cdot)$, is strictly decreasing in the number of concordant players until it reaches zero.*

Example 1. A simple dissonance function is:

$$d(s^i, n) = M(s^i) \cdot max(0, 1 - \frac{n}{\bar{N}}) \tag{2.3}$$

This function is at its maximum for a given action when $n = 0$ and this maximum, $M(s^i) > 0$, can be different for each action. Social dissonance decreases linearly in the number of concordant agents until $\bar{N}$. If at least $\bar{N}$ other agents have chosen the same action as agent i, he experiences no social dissonance.

2.4 Two-Action Games

Two-action games, where each agent has a choice set $S^i = \{A, B\}$, can be used to model many yes/no decisions: whether to follow a social norm, express support for a policy, or default on a loan.

---

[6] *$d(s^i, N-1)$ is the dissonance of choosing s^i when everyone else does and $d(s^i, 0)$ is the dissonance of being the only agent to choose action s^i.*

For each agent $i \in \{1, 2, ..., N\}$, let $\Delta u^i := u^i(B) - u^i(A)$. Arrange the agents in order of liking for B relative to A so that Δu^i is decreasing in i. There may be agents with identical values of Δu^i.[7]

2.4.1 Supermodularity

Games with strategic complementarities are also refered to as *supermodular* games. The key feature of these games is that agents' utility functions exhibit increasing differences: the relative benefit of choosing a "higher" action increases when the actions of other agents are higher. Before formally defining increasing differences, the action space should be ordered.

Define an ordering on the actions in $\{A, B\}$: let $B >_i A$ for each i.[8] Define the partial ordering $\geq$ over action profiles player-wise: for $s = (s^1, s^2, ..., s^N)$ and $s' = (s'^1, s'^2, ..., s'^N)$, $s \geq s'$ if $s^i \geq_i s'^i$ for all $i \in I$. [9]

Definition. An agent's payoff function U^i has **increasing differences** in s^i and s^{-i} if $s'^{-i} \geq s^{-i}$ implies

$$U^i(B, s'^{-i}) - U^i(A, s'^{-i}) \geq U^i(B, s^{-i}) - U^i(A, s^{-i}) \tag{2.4}$$

Proposition 1. *Any finite-agent two-action social dissonance game is supermodular.*

Proof. We need to show that condition (2.4) holds: each agent's payoff function has increasing differences in s^i and s^{-i}. Suppose that $s'^{-i} \geq s^{-i}$. If $s'^{-i} \geq s^{-i}$ then $N(s'^{-i}, B) \geq N(s^{-i}, B)$ and $N(s'^{-i}, A) \leq N(s^{-i}, A)$. Both $d(A, \cdot)$ and $d(B, \cdot)$ are nonincreasing in the second argument, so

[7] When there are some agents with identical Δu^i, it is useful to group them into types. Let there be $T \leq N$ types or distinct values of Δu^i for $i \in \{1, 2, ..., N\}$. Label the types $\{1, 2, ..., T\}$, let the numbers of agents of each type be $N_1, N_2, ..., N_T$ and let the differences in intrinsic utility for each type be $\Delta u^{t_1}, \Delta u^{t_2}, ..., \Delta u^{t_T}$.

[8] The ordering could also be $A >_i B$ and the game will still be supermodular. However, the increasing differences property will not hold under a player-specific ordering where $B >_i A$ for some players and $A >_i B$ for others. Keep in mind this is *not* a preference ordering.

[9] Define an ordering on an arbitrary list of player actions in the same way. For example, with $s^{-i} = (s^1, s^2, ..., s^{i-1}, s^{i+1}, ..., s^N)$ and $s'^{-i} = (s'^1, s'^2, ..., s'^{i-1}, s'^{i+1}, ..., s'^N)$, $s^{-i} \geq s'^{-i}$ if $s^j \geq_i s'^j$ for all $j \in \{1, 2, ..., N\}$ with $j \neq i$.

$$d(B, N(s^{-i}, B)) - d(B, N(s'^{-i}, B)) \geq 0$$

and

$$d(A, N(s'^{-i}, A)) - d(A, N(s^{-i}, A)) \geq 0$$

Combining these two inequalities yields

$$d(B, N(s^{-i}, B)) - d(B, N(s'^{-i}, B)) + d(A, N(s'^{-i}, A)) - d(A, N(s^{-i}, A)) \geq 0 \quad (2.5)$$

After rearranging (2.5) and adding $u^i(B) - u^i(A)$ to both sides, we obtain

$$u^i(B) - d(B, N(s'^{-i}, B)) - [u^i(A) - d(A, N(s'^{-i}, A))] \geq$$
$$u^i(B) - d(B, N(s^{-i}, B)) - [u^i(A) - d(A, N(s^{-i}, A))]. \quad (2.6)$$

This is equivalent to $U^i(B, s'^{-i}) - U^i(A, s'^{-i}) \geq U^i(B, s^{-i}) - U^i(A, s^{-i})$.

□

The supermodularity of two-action social dissonance games allows us to use the results of Milgrom and Roberts (1990) to derive properties of the equilibrium set. Existence of a pure-strategy equilibrium is guaranteed. Games which have a unique equilibrium are dominance-solvable. With respect to the ordering of action profiles described in Proposition 10, supermodularity guarantees the existence of a largest and smallest equilibrium, as well as monotonicity of the smallest and largest equilibria in response to changes in the parameters of the payoff functions. Monotone comparative statics results are discussed further in section 2.5.

2.4.2 Equilibria of Two-Action Games

Depending on the parameters $\{u^i\}_{i=0}^N$ and d, a social dissonance game may have different equilibrium sets. In some games, there are *monolithic* equilibria: those where every agent chooses the same action, regardless of intrinsic preferences. There may be equilibria where everyone feels free to act or express themselves in a way that reflects their intrinsic preferences. Such an equilibrium is called *true-to-type*.

Definition. A *true-to-type equilibrium* is an equilibrium action profile s such that for each i, if $\Delta u^i > 0$ then $s^i = B$ and if $\Delta u^i < 0$ then $s^i = A$.

Those agents with $\Delta u^i < 0$ are *type-a* agents and those with $\Delta u^i > 0$ are *type-b* agents. Agents who have $\Delta u^i = 0$ are indifferent, or *type-0* agents, and may play either A or B in a true-to-type equilibrium. Other equilibria are possible, but I first focus on identifying conditions for monolithic and true-to-type equilibria.

For a monolithic-A to be an equilibrium, the intrinsic utility gain to any agent from choosing B alone must not outweigh the additional social dissonance: $\Delta u^i \leq d(B, 0)$ for all $i \in \{1, 2, ..., N\}$. For monolithic-B to exist, $-\Delta u^i \leq d(A, 0)$ for all $i \in \{1, 2, ..., N\}$. Equivalently, we could just check these conditions for the extreme agents. Hence, for monolithic-A to exist, it must be that

$$\Delta u^1 \leq d(B, 0) \tag{2.7}$$

and for monolithic-B to exist,

$$-\Delta u^N \leq d(A, 0) \tag{2.8}$$

In true-to-type equilibrium, the incremental benefit of playing one's intrinsically preferred action outweighs any incremental social dissonance cost from doing so. Indifferent agents complicate matters in that they may be divided between actions A and B, but from each indifferent agent's perspective, the social dissonance of the other action must not be lower.

Let $N_a = \#\{i : \Delta u^i < 0\}$, $N_b = \#\{i : \Delta u^i > 0\}$, and $N_0 = N - N_a - N_b$. There is a true-to-type equilibrium if there exist non-negative integers N_0^A and N_0^B, with $N_0^A + N_0^B = N_0$, such that

$$-\Delta u^i \geq d(A, N_a + N_0^A - 1) - d(B, N_b + N_0^B) \ \forall i \text{ s.t. } \Delta u^i < 0 \tag{2.9}$$

$$\Delta u^i \geq d(B, N_b + N_0^B - 1) - d(A, N_a + N_0^A)\ \forall i \text{ s.t. } \Delta u^i > 0 \tag{2.10}$$

$$d(B, N_b + N_0^B) \geq d(A, N_a + N_0^A - 1) \text{ if } N_0^A > 0 \tag{2.11}$$

$$d(A, N_a + N_0^A) \geq d(B, N_b + N_0^B - 1) \text{ if } N_0^B > 0 \tag{2.12}$$

Conditions (2.9) and (2.10) can be verified by simply checking that (2.9) holds for the type-a with the least intense preference for A and (2.10) holds for the type-b with the least intense preference for B. Conditions (2.11) and (2.12) are ignored when there are no indifferent agents in the society.

Proposition 2 limits the set of equilibria by imposing a monotonicity requirement: If, in equilibrium, agent i chooses an action that he intrinsically likes less than agent j does, then agent j must also choose that action. The intuition behind the proof can be summarized as follows: If some type-a's choose B, it must be because the social dissonance of choosing B is lower than the social dissonance of choosing A. Otherwise, these type-a's would have chosen A, their intrinsically preferred action. Now, with the social dissonance of B being lower than A, type-b's are doubly glad to choose B: it is intrinsically preferred and offers lower social dissonance.

Proposition 2. *(Monotonicity of Equilibrium Action Profile): Let s be an equilibrium action profile. If for agents i and j, $u^j(s^i) - u^j(r^i) > u^i(s^i) - u^i(r^i)$ for $r^i \neq s^i$, then $s^j = s^i$.*

Proof. Suppose there is a player $i \in \{1, 2, ..., N\}$ who chooses action s^i in equilibrium action profile s and another player $j \in \{1, 2, ..., N\}$ such that $u^j(s^i) - u^j(r^i) > u^i(s^i) - u^i(r^i)$. s^i is a best response of i, so $u^i(s^i) - d(s^i, N(s^{-i}, s^i)) \geq u^i(r^i) - d(r^i, N(s^{-i}, r^i))$. It is easier to proceed by writing this best-response condition as

$$u^i(s^i) - u^i(r^i) \geq d(s^i, N(s^{-i}, s^i)) - d(r, N(s^{-i}, r^i))$$

Now, since player i is not counted in $N(s^{-i}, s^i)$ but will be counted in $N(s^{-j}, s^i)$, it must be that $N(s^{-j}, s^i) \geq N(s^{-i}, s^i)$. We don't know player j's action yet; it might be counted in $N(s^{-i}, r^i)$ but won't be counted in $N(s^{-j}, r^i)$, so $N(s^{-j}, r^i) \leq N(s^{-i}, r^i)$. Given only that d is non-increasing in its second argument, we know that $d(s^i, N(s^{-j}, s^i)) \leq d(s^i, N(s^{-i}, s^i))$ and $d(r^i, N(s^{-j}, r^i)) \geq d(r^i, N(s^{-i}, r^i))$. Combining these two inequalities, we have

$$d(s^i, N(s^{-i}, s^i)) - d(r^i, N(s^{-i}, r^i)) \geq d(s^i, N(s^{-j}, s^i)) - d(r^i, N(s^{-j}, r^i))$$

Stringing together j's stronger preference for s^i, i's best-response condition, and this last inequality concerning the social dissonance, we have

$$u^j(s^i) - u^j(r^i) > d(s^i, N(s^{-j}, s^i)) - d(r^i, N(s^{-j}, r^i))$$

which can be re-written to form j's best-response condition.
Only s^i is a best-response for j.

□

Monotonicity diminishes the set of action profiles that are candidates for equilibria. When all N agents have distinct Δu^i, there are $N+1$ candidates for equilibria, rather than 2^N, which is the number of possible action profiles without the monotonicity restriction. Monolithic-A and monolithic-B are possible equilibria, and can be checked with conditions (2.7) and (2.8). The other candidates for equilibria are any "intermediate" action profiles[10] where, with the agents in order of their preference for B from highest to lowest, $s^i = B$ for $i \leq k$ and $s^i = A$ for $i > k$. To test whether each of these action profiles is an equilibrium, we may just check the best-response conditions for the two marginal agents, k and $k+1$.

The discussion in the previous paragraph addressed games in which agents are organized in strictly descending order of their preference for B. There is the possibility that a group of agents have the same Δu^i, but some choose A and some choose B. *Asymmetric* equilibria have agents with the same intrinsic preferences choosing different actions. Monotonicity

[10] True-to-type is a special case.

rules out asymmetric behavior within more than one group of agents of the same Δu^i (see Lemma 24 in the appendix). However, as Proposition 3 shows, if Assumption 3 holds, there can be no asymmetric behavior in equilibrium within a group of agents with identical Δu^i who are not intrinsically indifferent.

Proposition 3. *Suppose Assumption 3 holds. Let s be an equilibrium action profile. If for agents i and j, $\Delta u^i \neq 0$, and $u^j(s^i) - u^j(r^i) \geq u^i(s^i) - u^i(r^i)$ for $r^i \neq s^i$, then $s^j = s^i$.*

Proof. Proposition 2 states that, even without Assumption 3, if in equilibrium, $u^j(s^i) - u^j(r^i) > u^i(s^i) - u^i(r^i)$, then $s^j = s^i$. I will prove by contradiction that for two agents i and j with $u^j(B) - u^j(A) = u^i(B) - u^i(A) \neq 0$, in equilibrium s, it must be that $s^j = s^i$. Suppose to the contrary there is an equilibrium s and there are two agents i and j such that $\Delta u^i = \Delta u^j \neq 0$, $s^i = A$, and $s^j = B$.

A is a best response to s^{-i}, and B is a best response to s^{-j}. By definition of best response,

$$\begin{aligned} u^i(A) - d(A, N(s^{-i}, A)) &\geq u^i(B) - d(B, N(s^{-i}, B)) \\ u^j(A) - d(A, N(s^{-j}, A)) &\leq u^j(B) - d(B, N(s^{-j}, B)) \end{aligned} \tag{2.13}$$

Rearranging these, we get

$$\begin{aligned} u^i(A) - u^i(B) &\geq d(A, N(s^{-i}, A)) - d(B, N(s^{-i}, B)) \\ u^j(A) - u^j(B) &\leq d(A, N(s^{-j}, A)) - d(B, N(s^{-j}, B)) \end{aligned}$$

Since $\Delta u^j = \Delta u^i$,

$$d(A, N(s^{-j}, A)) - d(B, N(s^{-j}, B)) \geq d(A, N(s^{-i}, A)) - d(B, N(s^{-i}, B))$$

which is to say

$$d(A, N(s^{-j}, A)) - d(A, N(s^{-i}, A)) + [d(B, N(s^{-i}, B)) - d(B, N(s^{-j}, B))] \geq 0 \tag{2.14}$$

$N(s^{-i}, A)$, the count of other agents from the perspective of i does not include i, but $N(s^{-j}, A)$ does. Since agent j chose B, neither $N(s^{-i}, A)$ nor $N(s^{-j}, A)$ include j. Hence,

$N(s^{-j}, A) > N(s^{-i}, A)$. Likewise, $N(s^{-i}, B) > N(s^{-j}, B)$. These last two inequalities imply

$$d(A, N(s^{-j}, A)) - d(A, N(s^{-i}, A)) \leq 0$$

$$d(B, N(s^{-i}, B)) - d(B, N(s^{-j}, B)) \leq 0$$

due to the non-decreasing nature of the social dissonance functions.

Under Assumption 3, either these inequalities are strict, which contradicts (2.14), or the social dissonance is zero. If social dissonance is zero, (2.13) implies $u^i(A) \geq u^i(B)$ and $u^j(A) \leq u^j(B)$. Since $u^i(A) - u^i(B) = u^j(A) - u^j(B)$, it must be that both equal zero. However, this contradicts the assumption that $\Delta u^i \neq 0$ and $\Delta u^j \neq 0$. □

Example 2 in 2.6.1 illustrates that an asymmetric equilibrium may exist without Assumption 3, and the discussion in 2.6.2 addresses asymmetric behavior within a group of intrinsically indifferent agents. The multiplicity of equilibria due to asymmetry among indifferent agents suggests the potential for a coordination problem with welfare implications. The Dixit game in 2.6.3 is an example of a game where every agent has different intrinsic preferences.

2.5 Comparative Statics

Action profiles are ordered according to the partial order $\geq$ defined before Proposition 1. The lowest action profile is monolithic-A, and the highest is monolithic-B. Under Assumption 3, the set of (pure-strategy) equilibria of a social dissonance game can be any non-empty subset of the monotone action profiles. An *increase in an equilibrium set* is an increase, with respect to the order $\geq$, of either the lowest equilibrium, the highest equilibrium, or both. In other words, an increase is a shift toward profiles where more agents play action B. A *decrease in an equilibrium set* is defined conversely as a shift toward profiles where more agents play A.

Define the vector $\Delta u := (\Delta u^i)_{i=1}^N$ and let $\Delta d := (\Delta d_n)_{n=0}^{N-1}$ be a vector of differences in social dissonance between action A and action B faced by an agent when n other agents choose B. That is, $\Delta d_n = d(A, N-1-n) - d(B, n)$ for $n \in \{0, 1, ..., N-1\}$. Recall from observing the best response condition (2.2) that the equilibrium set depends on Δu and Δd. The effects of changes in these parameters are verified separately. Proposition 4 shows that a shift in the intrinsic utility vector, Δu, toward one action will be accompanied by a shift in the equilibrium set toward the same action. Proposition 6 shows that the equilibrium set changes monotonically when there is an overall change in the social dissonance for one action relative to the other.

Proposition 4. *The equilibrium set of a social dissonance game is non-decreasing in Δu.*

Proof. Theorem 6 and its corollary in Milgrom and Roberts (1990) state that a family of supermodular games, such as two-action social dissonance games, with payoff functions that have an additional increasing differences property with respect to some exogenous parameter (Δu), has equilibrium sets that are non-decreasing in Δu. This increasing differences property is described and verified in Lemma 5. □

Lemma 5. *For all i, U^i has increasing differences in s^i and Δu for any fixed s^{-i}.*

Proof. One only must verify that, for any fixed s^{-i}, the difference $U^i(B, s^{-i}) - U^i(A, s^{-i})$ is increasing in Δu.

$$\begin{aligned} U^i(B, s^{-i}) - U^i(A, s^{-i}) &= u^i(B) - d(B, N(s^{-i}, B)) - (u^i(A) - d(A, N(s^{-i}, A))) \\ &= (u^i(B) - (u^i(A)) - (d(B, N(s^{-i}, B)) - d(A, N(s^{-i}, A))) \\ &= \Delta u^i + (d(A, N(s^{-i}, A)) - d(B, N(s^{-i}, B))) \end{aligned}$$

An increase in Δu^i clearly increases $U^i(B, s^{-i}) - U^i(A, s^{-i})$, while any increase in Δu^j for $j \neq i$ has no effect. Hence, $U^i(B, s^{-i}) - U^i(A, s^{-i})$ is increasing in Δu.

□

Proposition 6. *The equilibrium set of a social dissonance game is non-decreasing in* Δd*.*

Proof. As with Proposition 4, the result follows from Theorem 6 in Milgrom and Roberts after verifying that U^i has increasing differences in s^i and Δd for any fixed s^{-i}. The analog to the equation in Lemma 5 is

$$\begin{aligned} U^i(B, s^{-i}) - U^i(A, s^{-i}) &= u^i(B) - d(B, s^{-i}) - (u^i(A) - d(A, s^{-i})) \\ &= u^i(B) - u^i(A) + (d(A, s^{-i}) - d(B, s^{-i})) \\ &= u^i(B) - u^i(A) + (d(A, N(s^{-i}, A)) - d(B, N(s^{-i}, B))) \\ &= u^i(B) - u^i(A) + (d(A, N - 1 - N(s^{-i}, B)) - d(B, N(s^{-i}, B))) \\ &= u^i(B) - u^i(A) + \Delta d(N(s^{-i}, B)) \end{aligned}$$

□

The equilibrium set varies in a nondecreasing manner with any increase in Δd. This increase in Δd could be any change in the dissonance function from d to d' where $d'(A, n) \geq d(A, n)$ for all $n \in \{0, 1, ..., N-1\}$ and $d'(B, n) \leq d(B, n)$ for all $n \in \{0, 1, ..., N-1\}$, or any change from d to d' that makes A worse relative to B in terms of social dissonance for a fixed action profile of other agents.

2.6 Example Games

The examples that follow are included to provide a demonstration of the general results on the characterization of equilibrium and comparative statics. They may be used to quickly model common situations in which social dissonance is a factor. Because these examples are intended to stand alone and provide ready-made solutions, some of the results may seem repetitive of those from earlier sections.

2.6.1 Game with Two Types

Imagine an issue on which society is divided into two camps of agents with identical preferences, all of whom are forced to choose between A and B. In this society, N_a of the agents are of type-a and intrinsically like action A better, while N_b, or $N - N_a$, of the agents are of type-b and intrinsically like action B better. The intrinsic utility functions for the two types, u^a and u^b, are such that $u^a(A) > u^a(B)$ and $u^b(B) > u^b(A)$. Let $\Delta u^a := u^a(A) - u^a(B)$ and $\Delta u^b := u^b(B) - u^b(A)$ be the gaps in intrinsic utility between the preferred action and non-preferred action for type-a and type-b agents. The game is thus defined by $\{N_a, N_b, \Delta u^a, \Delta u^b, d\}$.

Monolithic-A is an equilibrium if and only if $\Delta u^b \leq d(B, 0)$ and monolithic-B is an equilibrium if and only if $\Delta u^a \leq d(A, 0)$. There exists a true-to-type equilibrium if

$$\Delta u^a \geq d(A, N_a - 1) - d(B, N_b) \tag{2.15}$$

and

$$\Delta u^b \geq d(B, N_b - 1) - d(A, N_a) \tag{2.16}$$

Due to the supermodularity of the game, it has at least one pure-strategy equilibrium for arbitrary parameters. It is also possible to prove the existence of a pure-strategy equilibrium by verifying that either the true-to-type conditions or at least one of the monolithic equilibrium conditions will always hold. This proof is outlined in Appendix A.

From Proposition 2 it follows that if, in equilibrium some agents of one type play against-type, then all agents of the other type play true-to-type. Precluded as equilibria are any action profiles where both types choose against type, there is asymmetry within both types, or one type behaves asymmetrically and the other chooses against type.

When Assumption 3 holds, the set of candidates for equilibria is limited to monolithic-A, true-to-type, and monolithic-B. The equilibrium set can be any non-empty subset of the three candidate equilibria, depending on the parameters of the game.

This raises the question of whether members of the same group may choose different actions in equilibrium. An asymmetric equilibrium can exist if Assumption 3 does not hold, as the following example illustrates.

Example 2. Suppose a society has $N_a = 5$ and $N_b = 5$. Suppose $\Delta u^a = 0.5$, $\Delta u^b = 2$ and the social dissonance function is decreasing, but not strictly. d is such that $d(\cdot, 0) = 1$ and if $2 \leq n \leq 4$, $d(A, n) = 0.75$ and $d(B, N-n-1) = 0.25$. Type-b's have a dominant strategy of choosing B. If there is asymmtery within type-a, with 3 choosing A and 2 choosing B, a type-a agent who chooses A sees 2 others choosing A and 7 others choosing B. A type-a who chooses B sees 3 others choosing A and 6 others choosing B. The payoffs to the type-a's are $0.5 - 0.75$ for choosing A and $0 - 0.25$ for B. There is no incentive for any individual to choose otherwise; the action profile described is an asymmetric equilibrium.

When a game has a unique monolithic equilibrium, the players who are choosing against their intrinsic preferences can be said to be "entrapped." Conditions for this are generalized in section 2.7, but the idea is that one type of player has strong enough preferences to overcome any possible social dissonance, and the other type does not represent enough of the population to reduce social dissonance to a comfortable level. In games with a unique true-to-type equilibrium, none of the agents care enough about social dissonance for it to have any effect on their decisions.

Games that have multiple equilibria, especially those with all three, suggest the need for welfare analysis and a coordination device to ensure that agents settle on the best equilibrium. For example, if true-to-type is an equilibrium and it Pareto-dominates[11] the two monolithic equilibria, then a government may wish to signal to agents in a way that would induce them to choose according to intrinsic preferences.

In the notation of the comparative statics result, Proposition 4, changes in the population are a shift in Δu^i. In the two-type game, replacing a type-a player i with a type-b player is the same as increasing Δu^i from $-\Delta u^a$ to Δu^b. Increasing Δu^b or decreasing

[11] Action profile s *Pareto dominates* s' if $U^i(s) \geq U^i(s') \forall i$ and $U^i(s) > U^i(s')$ for at least one i.

Δu^a would also be represented as increases in Δu^i. All of these changes would lead to non-decreasing shifts[12] in the equilibrium set.

Non-monotonic changes in the dissonance function can affect the equilibrium set. In the following example, a change from d to d' where $d' \geq d$, removes true-to-type from the equilibrium set.

Example 3. Suppose in the two-type game, $N_a = 15$ and $N_b = 5$. Let $\Delta u^a = \Delta u^b = 0.5$ and let $d(s^i, n) = max(0, 1 - \frac{n}{6})$. The equilibria are monolithic-A, monolithic-B, and true-to-type. Now, if instead, the dissonance function is $d'(s^i, n) = 2 \cdot max(0, 1 - \frac{n}{6})$, only monolithic-A and monolithic-B are equilibria.

Increasing the entire social dissonance function in a multiplicative manner is representative of an institution that discourages diversity and nonconformity in general. In societies like the one in Example 3 where the population is unbalanced, such a change in the institution is more likely to be eliminate true-to-type from the equilibrium set.

2.6.2 Game with Two Types and Indifferent Agents

Elections and opinion polls on important issues often force citizens to choose between one of two alternatives. Voters are allowed to abstain and poll respondents are allowed to say they are indifferent, but a citizen who is intrinsically indifferent may choose one way or another due to some perceived pressure to make a clear choice. It is also possible that the citizen's mood, the ordering of the alternatives, or some other random factor can sway his vote to either side of the fence. The opinions and votes of these indifferent citizens can change the course of history.

To investigate the effect of forced decisions by indifferent agents on the payoffs to others, include a third type of agent, the type-0. These agents decide solely on the basis of social dissonance concerns. Suppose there are N_a agents of type-a, N_b type-b agents, and N_0 type-0 agents. The game is described by $(N_a, N_b, N_0, \Delta u^a, \Delta u^b, d)$.

[12] The order of equilibria, from lowest to highest, is monolithic-A, true-to-type, monolithic-B.

Proposition 2 restricts the set of action profiles that may be equilibria. In particular, if there is an agent of type-0 who chooses A in equilibrium, than all type-a agents must choose A; if there is a type-0 agent who chooses B, then all type-b's choose B. [13]

In order to have a true-to-type equilibrium, there must exist non-negative integers N_0^A and N_0^B, with $N_0^A + N_0^B = N_0$, such that

$$\Delta u^a \geq d(A, N_a + N_0^A - 1) - d(B, N_b + N_0^B) \tag{2.17}$$

$$\Delta u^b \geq d(B, N_b + N_0^B - 1) - d(A, N_a + N_0^A) \tag{2.18}$$

$$d(B, N_b + N_0^B) \geq d(A, N_a + N_0^A - 1)\} \text{ if } N_0^A > 0 \tag{2.19}$$

$$d(A, N_a + N_0^A) \geq d(B, N_b + N_0^B - 1)\} \text{ if } N_0^B > 0 \tag{2.20}$$

As in the two-type game, if $\Delta u^b \leq d(B, 0)$, monolithic-A is an equilibrium and if $\Delta u^a \leq d(A, 0)$, monolithic-B is an equilibrium.

In equilibrium, there can be asymmetric behavior among type-0 agents. If Assumption 3 is imposed, we can limit the set of asymmetric equilibria to consider. If type-a and type-b both choose true-to-type, type-0 choices in equilibrium can be asymmetric, but this requires that social dissonance is zero for both actions.

Proposition 7. *Under Assumption 3, if action profile s is an equilibrium where type-0 agents behave asymmetrically, then under s, social dissonance must be zero for all agents.*

[13] If type-b chooses A in equilibrium, then that equilibrium is monolithic-A. If type-a chooses B, that equilibrium is monolithic-B.

Proof. Suppose in equilibrium there is asymmetric behavior within the group of type-0 agents, with N_0^A choosing A and N_0^B choosing B. The N_a type-a's choose A and the N_b type-b's choose B. The condition for a type-0 to choose A is $d(A, N_a + N_0^A - 1) \leq d(B, N_b + N_0^B)$. The condition for a type-0 to choose B is $d(B, N_b + N_0^B - 1) \leq d(A, N_a + N_0^A)$. By monotonicity of d, $d(A, N_a + N_0^A - 1) \geq d(A, N_a + N_0^A)$ and $d(B, N_b + N_0^B - 1) \geq d(B, N_b + N_0^B)$. Stringing these together, we have

$$d(B, N_b+N_0^B) \geq d(A, N_a+N_0^A-1) \geq d(A, N_a+N_0^A) \geq d(B, N_b+N_0^B-1) \geq d(B, N_b+N_0^B)$$

which implies that all of these are equal. Since d is strictly decreasing until it reaches zero, $d(A, N_a + N_0^A - 1) = d(A, N_a + N_0^A)$ implies that dissonance must be zero.

□

With type-0 agents in a society, finding all equilibria is more complicated than just checking the list of sufficient conditions starting with (2.17). First, check to see if the conditions for monolithic equilibria are satisfied. Next, use the true-to-type conditions to test for the two equilibria where type-a plays A, type-b plays B, and type-0's either all play A or all play B. Finally, use proposition 7 to find the asymmetric equilibria with the minimum and maximum number of type-0's playing A.

Example 4. Suppose $N_a = N_b = N_0 = 10$. Let $\Delta u^a = \Delta u^b = 0.5$ and let $d(s^i, n) = max(0, 1 - \frac{n}{12})$. There are monolithic-A and monolithic-B equilibria, since $\Delta u^b \leq d(B, 0)$ and $\Delta u^a \leq d(A, 0)$. There are two true-to-type equilibria where all type-0's coordinate on either action. Finally, there are true-to-type equilibria where type-0's behave asymmetrically. By Proposition 7, social dissonance must be zero for both actions, so there must be at least three type-0 agents playing either action.

The society described in the example above is one where all of the true-to-type equilibria Pareto-dominate the monolithic equilibria. Note also that in Example 4, the equilibria where type-0 agents behave asymmetrically all Pareto-dominate the equilibria where type-0's all choose the same action. This suggests that society would benefit from a coordination device so the indifferent agents divide into two groups: some who choose A and some who choose B.

2.6.3 Dixit Game

N players must each decide whether to maintain status quo (action A) or join a club (action B). Joining the “club” can represent adoption of a new technology or standard, or the holding of a progressive opinion on an issue. Suppose the gap in intrinsic utility between B and A decreases linearly with player type, i, which runs from 1 to N. Set $u^i(B) = \alpha - \beta i$ and $u^i(A) = 0$, so $\Delta u^i := u^i(B) - u^i(A)$ is equal to $\alpha - \beta i$. Suppose the social dissonance is also linear. The social dissonance function is

$$\begin{aligned} d(A, n) &= \gamma(N - 1 - n) \\ d(B, n) &= \tau(N - 1 - n) \end{aligned} \tag{2.21}$$

The intrinsic utility functions and the dissonance functions together give the payoffs:

$$\begin{aligned} &U^i(A, s^{-i}) = -\gamma(N - 1 - N(s^{-i}, A)) \\ &U^i(B, s^{-i}) = \alpha - \beta i - \tau(N - 1 - N(s^{-i}, B)) \\ &\beta, \gamma, \tau > 0 \end{aligned} \tag{2.22}$$

By condition (2.7), monolithic-A is an equilibrium if $\alpha - \beta \leq \tau(N - 1)$. By condition (2.8), monolithic-B is an equilibrium if $\beta N - \alpha \leq \gamma(N - 1)$. Other equilibria are possible, and will obey the monotonicity property described in Proposition 2.

Comparative statics follow from the results of section 2.5. An increase in α will lead to a shift in the equilibrium set toward action profiles where more agents play B, while an increase in β will shift the equilibrium set toward action A. Increasing γ or decreasing τ increases the social dissonance of A relative to B and will lead to a shift in the equilibrium set toward profiles where B is played by more agents.

2.7 Cascades and Entrapment

In some social dissonance games, the choices of the agents who have the most intense intrinsic preferences set off a cascade of externalities that lead other agents to favor joining them. This cascade continues until all agents are convinced to join, due to the increasing social dissonance cost of choosing otherwise.

A **cascade equilibrium** is a monolithic equilibrium in a dominance-solvable game. It follows from the definition of dominance solvability that a cascade equilibrium is unique. In supermodular games, uniqueness of a pure strategy equilibrium guarantees dominance solvability (Milgrom and Roberts, 1990). However, uniqueness alone does not define a cascade equilibrium; a true-to-type equilibrium may be the unique solution from iterated elimination of strongly dominated actions.

I discuss conditions for establishing monolithic-B as a cascade equilibrium, but this can easily be applied to the case where monolithic-A is the only equilibrium. Recall that monolithic-B is an equilibrium if no one gains enough intrinsic utility from A to overcome the dissonance of choosing it alone: $\forall i, d(A, 0) > -\Delta u^i$.

In order to preclude the monolithic-A equilibrium, there must be an agent for whom B is a strictly dominant strategy. That is, for some agent i it must be that $u^i(B) - d(B, 0) > u^i(A) - d(A, N-1)$. Since $d(A, N-1) = 0$ and the players are arranged in decreasing order of preference for A, this is

$$\Delta u^1 > d(B, 0) \tag{2.23}$$

There must also be a similar condition for the rest of the agents that leads to the cascade into monolithic-B. Knowing that there is a agent ($i = 1$) that will always choose B, the agent with the next-strongest preference for B must find B to be a dominant strategy. The cascade of iterated elimination of strictly dominated actions continues until all agents, including those who most strongly prefer A, are convinced to play B to avoid the social dissonance of choosing A. This second condition is $\forall i \in \{2, 3, ..., N\}$,

$$\Delta u^i > d(B, i-1) - d(A, N-1-i) \tag{2.24}$$

From these conditions follow the sufficient conditions for monolithic-B to be a cascade equilibrium established in Dixit (2003) for the game in section 2.6.3: (1) there exists a player whose preference for B is strong enough to overcome the social dissonance of being the only member of the club. (2) The marginal player-to-player decrease in strength of preference for B is less the marginal reduction in social dissonance of choosing B when an additional player joins. [14] Dixit's second condition is stronger than necessary, because it is the sum of marginal dissonance that matters in condition (2.24).

Example 5. Suppose there are four agents in a society and the social dissonance function is $d(s^i, n) = 1 - \dfrac{n}{N-1}$. One agent has a preference for B that is intense enough to overcome any social dissonance concerns ($\Delta u^1 = 1.5$), two agents have a less intense preference for B ($\Delta u^2 = \Delta u^3 = 0.75$), and one agent prefers A ($\Delta u^4 = -0.75$). In this society, monolithic-B is a cascade equilibrium.

The sufficient conditions (2.23) and (2.24) for monolithic-B to be a cascade equilibrium in the static simultaneous-move game are also sufficient for a subgame perfect equilibrium in a sequential version of the game to result in monolithic-B. [15] To see why this is true, consider an extensive form game with N players, [16] each of whom moves exactly once in an arbitrary order. Information is perfect and complete.

Supposing condition (2.23) holds, player 1 will always choose B. That is, paths in the game tree in which player 1 chooses A will not be played in equilibrium. If player 1 moves last, then player 2 will by backward induction, realize that B will be chosen, and due to (2.24) only B as the best response. If player 1 moves before player 2, then player 2 again

[14] Consistent with the specification in 2.6.3, this is (1)$\alpha - \beta i > \tau(N-1)$ and (2) $\beta < \gamma + \tau$.
[15] Dixit (2003) has a similar result.
[16] This can be generalized to a game with replicants.

chooses B as a best response. Players 3 through N will in turn be subject to the same logic, and cascade into monolithic-B follows as the only equilibrium outcome.

2.7.1 Entrapment

There are equilibria in which some agents would be better off if others chose differently, but are "entrapped" into choosing an action they do not intrinsically prefer.[17] In Example 5, monolithic-B is an entrapment for agent 4. This concept of entrapment can be generalized to include any action profile where at least one agent is negatively affected by the social dissonance of others' choices. Even a true-to-type equilibrium may be an entrapment, if it is not the best outcome for some of the players.

Definition. Action profile s is an entrapment for player i if there is another action profile $\hat{s}$ such that $U^i(\hat{s}) > U^i(s)$.

Finding such an $\hat{s}$ that is a candidate for improvement is easy: choose the profile that is monolithic in i's intrinsically preferred action. This will maximize i's payoff since u^i is maximized and d evaluates to zero.

Lemma 8. *For any non-true-to-type action profile, there is an agent i for whom a monolithic profile of i's intrinsically preferred action is strictly better.*

Proof. Suppose s is not a true-to-type action profile. There is an agent i for whom $u^i(s^i) < u^i(r^i)$ where r^i is the other action in i's choice set. Let $\hat{s}$ be the monolithic action profile where for all j, $\hat{s}^j = r^i$. $d(r^i, N(\hat{s}^{-i}, r^i)) = 0$ and $d(s^i, N(s^{-i}, s^i)) \geq 0$. $u^i(r^i) - d(r^i, N(\hat{s}^{-i}, r)) > u^i(s^i) - d(s^i, N(s^{-i}, s^i))$, which is to say $U^i(\hat{s}) > U^i(s)$. □

Monolithic-B is an entrapment for agents with $\Delta u^i < 0$ but is optimal for agents with $\Delta u^i > 0$. In order for this to be the result of a cascade (and satisfy the cascade conditions),

[17] Dixit(2003) describes entrapment as "the result that everyone joins a club which several members may continue to dislike."

the dissonance of choosing B as part of a minority must be less than the dissonance of choosing A as a member of the majority.

When the dissonance associated with action A is large relative to the dissonance of B, a small group of people with intense preferences for B can pressure a majority with a less intense preference for A. Examining the cascade condition again, if more than half of society is entrapped in the monolithic-B cascade equilibrium, there is a agent $i < N/2$ for whom $\Delta u^i > d(B, i-1) - d(A, N-i)$.

2.8 Games with More than Two Actions

Games with two actions allow us to study choices in many settings, but in others, there are more than two actions available. In opinion polls and voting, citizens may be faced with a choice between two salient candidates or viewpoints, and a third option to abstain from expressing an opinion or casting a vote.

In general, social dissonance games with more than two actions are not supermodular. To see this, consider a game with three actions, $\{A, B, C\}$ with an arbitrary ordering $A > B > C$. If, as before, social dissonance results from *any* discordant action, there may or may not be an incremental benefit to raising one's action from C. If a group of other agents increase their actions from B to A, this reduces the dissonance and thereby increases the incremental benefit of increasing one's action from C to A. However, an agent considering an increase from C to B may now face a greater level of dissonance and, hence, a lower incremental benefit.

Games that are supermodular are guaranteed to have pure strategy equilibria and monotone comparative statics properties. Games which are not supermodular may still have pure strategy equilibria, but additional analysis would be needed to identify them. The examples that follow are supermodular three-action games.

Example 6. Consider a situation where citizens are asked for their opinions on an election or actually vote, knowing that they may be asked later about their choices.[18] Suppose now that those who choose one of two frontrunners will face dissonance from those who do not, and those who choose a third-party candidate or otherwise abstain from choosing one of the frontrunners will experience no social dissonance.[19]

The dissonance functions for A, M, and B follow the form:

$$\begin{aligned} d(A, s^{-i}) &= d(A, N(s^{-i}, A)) \\ d(M, s^{-i}) &= 0 \\ d(B, s^{-i}) &= d(B, N(s^{-i}, B)) \end{aligned} \tag{2.25}$$

where $d(A, \cdot)$ and $d(B, \cdot)$ are decreasing in their arguments.

If $s'^{-i} > s^{-i}$, then $N(s'^{-i}, A) \leq N(s^{-i}, A)$ and $N(s'^{-i}, B) \geq N(s^{-i}, B)$. It follows that $d(A, s'^{-i}) \geq d(A, s^{-i})$ and $d(B, s^{-i}) \leq d(B, s'^{-i})$. As specified, $d(M, s'^{-i}) = d(M, s^{-i})$. There are three basic categories of increases in the actions of others: a group of agents can shift from A to B, from A to M, or from B to M. Isolating a single category of others' increase in action allows us to separately consider its effect on the incremental benefit of a single agent's increase in action. In order for the increasing differences property to hold in general, any increase in an agent's own action must be more beneficial when there is any category of increase in others' actions. The differences in the incremental benefit for each

[18] DellaVigna, List, Malmendier and Rao (2013) find there is a cost associated with both not voting and lying about one's choice after voting.

[19] The description can be generalized to assign a *constant* level of dissonance to those who choose neither of the frontrunners. Having zero social dissonance is consistent with Assumption 5.

possible increase in an agent's own action are

$$
\begin{aligned}
U^i(M,s'^{-i}) - U^i(A,s'^{-i}) - &[U^i(M,s^{-i}) - U^i(A,s^{-i})] \\
&= d(M,s^{-i}) - d(M,s'^{-i}) + d(A,s'^{-i}) - d(A,s^{-i}) \\
U^i(B,s'^{-i}) - U^i(M,s'^{-i}) - &[U^i(B,s^{-i}) - U^i(M,s^{-i})] \\
&= d(B,s^{-i}) - d(B,s'^{-i}) + d(M,s'^{-i}) - d(M,s^{-i}) \\
U^i(B,s'^{-i}) - U^i(A,s'^{-i}) - &[U^i(B,s^{-i}) - U^i(A,s^{-i})] \\
&= d(B,s^{-i}) - d(B,s'^{-i}) + d(A,s'^{-i}) - d(A,s^{-i})
\end{aligned}
\tag{2.26}
$$

It follows from the specification in (2.25) that $d(M,s^{-i}) - d(M,s'^{-i}) = 0$; thus, the expressions in (2.26) are non-negative. Hence, under the ordering of actions, $B > M > A$, with a dissonance function following the form of (2.25), the payoff functions U^i exhibit increasing differences in s^i and s^{-i} and the game is supermodular.[20]

Example 7. Suppose that in the example above that citizens choosing one of the frontrunners only experience social dissonance from people choosing the other frontrunner.

$$
\begin{aligned}
d(A,s^{-i}) &= d(A, N(s^{-i},A) + N(s^{-i},M)) \\
d(M,s^{-i}) &= 0 \\
d(B,s^{-i}) &= d(B, N(s^{-i},B) + N(s^{-i},M))
\end{aligned}
\tag{2.27}
$$

where $d(A,\cdot)$ and $d(B,\cdot)$ are decreasing in their arguments.

Payoffs in three-action games with this type of dissonance function also have increasing differences in own and others' actions. Therefore, such games are supermodular.

Example 8. Consider an election with three candidates or an issue with three leading perspectives, A, M, and B. Citizens who choose A face social dissonance from those who

[20] Supermodularity also holds under the ordering $A > M > B$.

choose otherwise, but citizens who choose M or B *tolerate* one another and only face dissonance from those who choose A.

$$
\begin{aligned}
d(A, s^{-i}) &= d(A, N(s^{-i}, A) \\
d(M, s^{-i}) &= d(M, N(s^{-i}, B) + N(s^{-i}, M)) \\
d(B, s^{-i}) &= d(B, N(s^{-i}, B) + N(s^{-i}, M))
\end{aligned}
\tag{2.28}
$$

where $d(A, \cdot)$, $d(M, \cdot)$, and $d(B, \cdot)$ are decreasing in their arguments. If we add to the specification that $d(M, \cdot) = d(B, \cdot)$, then the game is supermodular.

2.9 Conclusion

Social dissonance games provide a framework for examining the effects of institutions and individuals' intrinsic preferences in situations where the conformity motive is present. Equilibria must obey a monotonicity property: if one agent chooses an action, any agent who intrinsically likes the action more will also choose it. Two-action social dissonance games are supermodular. As such, we can apply the results of Milgrom and Roberts (1990) to generate clear qualitative predictions of how the equilibrium set will change. More specific information about agents' intrinsic values or the functional form of the dissonance function[21] allow for identification of the equilibrium set.

There can be multiple equilibria in a social dissonance game. Without a clear signal about the intentions of others, individuals in a society may fail to coordinate on the choices they intrinsically prefer. A monolithic equilibrium may be good for one subgroup, but the true-to-type equilibrium is not necessarily worse for them if social dissonance is eliminated at an agreement level well below 100%. In such cases, encouraging true-to-type behavior maximizes overall welfare. Managing the behavior of indifferent agents can help to avoid a suboptimal outcome[22] and entrapment of those holding a minority opinion. In the case

[21] See my Example 2.3.
[22] See Example 4 in section 2.6.2.

of aggressive changes in institutions, reducing social dissonance for marginalized citizens may also help to improve overall welfare.

Not all instances of social dissonance are counterproductive to progress. Governments can use social dissonance as a lever to change social norms. An otherwise expensive incentive program can be made more affordable by paying only some individuals to change behavior, then allowing the conformity motive to influence others to change. If, on the other hand, social contagion starts to become a problem, an institution may respond by raising social dissonance associated with the contagious action.

The model in this paper can be extended to dynamic games with imperfect information. We should note that misinformation and its distortion of social dissonance can slow down the process of a change in a social norm. A 2013 Gallup poll found that despite majority support for the legalization of same-sex marriage, 63% of respondents believed that the majority opposed legalization. Transparency of opinions will likely lead to a further shift in opinions to the point where a greater percentage of the population supports the legalization of same-sex marriage.

Chapter 3

Social Dissonance in a Model with a Continuum of Agents

3.1 Introduction

Social dissonance games can be used to model situations where the conformity motive affects individual decision-making. This includes responses to opinion polls, social behavior, and choices in markets. In a previous essay, I defined static social dissonance games with a finite number of agents and characterized the equilibria that could arise in such games with two actions. The finite-agent model is fitting when there are a small number of agents who are all aware of each other's actions, but when the population becomes larger and agents are only aware of actions on the aggregate level, a model with a continuum of agents in the style developed by Aumann (1964) is more appropriate.

In this paper, I define social dissonance games with a continuum of agents and characterize their equilibrium sets. Equilibrium action profiles are monotone in the sense that the group of players who choose a particular action are those with the strongest intrinsic preference for that action. In a continuum-of-agents model, each individual agent has no effect on the behavior of other agents. As a consequence of this, asymmetric behavior may be part of an equilibrium, even within a group of agents with a clear intrinsic preference for one action over the other.[1] A group of "sell-outs" going against their intrinsic preferences, with an inability to make a coordinated change, can hold in place a norm they do not favor.

Two-action social dissonance games belong to a larger class, supermodular games, which have been studied by Topkis (1979), Vives (1990), Milgrom and Roberts (1990), and Yang and Qi (2013). They are also a subclass of aggregative games, as studied by Acemoglu and Jensen (2010). In two-action social dissonance games, the set of equilibria depends only on each player's *difference* in utility between the two actions, and the *difference* in dissonance between the two actions for each aggregate action profile. I use the results of Yang and Qi to establish monotone comparative statics for equilibrium sets with respect to changes in intrinsic utilities and dissonance functions.

The framework of social dissonance games can be used in empirical studies of how changes in institutions - messages from media, lawmakers, and leaders - affect aggregate behavior. I provide some examples of how to compute equilibria for two-action games with dissonance functions and distributions of intrinsic utilities that are "manageable." For instance, if dissonance is linearly decreasing in level of agreement and intrinsic utility is linearly distributed across the population, there is a single equilibrium, described by a threshold agent. This threshold agent corresponds to a split in the behavior of the population into two groups that each choose one of the two actions.

This essay is organized as follows: In section 3.2, I describe the general class of social dissonance games with a continuum of agents. Section 3.3 begins the study of two-action games, and section 3.4 discusses some examples and the computation of their equilibria. Section 3.5 introduces games in which dissonance decreases linearly, and section 3.6 suggests uses for this specification in an empirical setting. Section 3.7 discusses dynamic games. Section 3.8 concludes.

[1] An asymmetric action profile is one in which agents with the same intrinic utilities for the actions choose different pure strategies.

3.2 The Model

The set of agents, I, is represented by the closed unit interval $[0, 1]$, endowed with Lebesgue measure μ. The agents have a common finite action set C and the joint action space S is the set of measurable functions from $[0, 1]$ to C. For any action profile $s \in S$, the action of player i given by $s(i)$. The measure of agents choosing action $c \in C$ in profile s is $\mu_s(c) := \mu\{i : s(i) = c\}$.[2]

Each agent $i \in [0, 1]$ has an intrinsic utility function $u^i : C \to \mathbb{R}$ and all agents share the same social dissonance function $d : C \times [0, 1] \to \mathbb{R}$ which takes as arguments a player's own action and the measure of agents choosing that action. The payoff to agent i, U^i, is determined by the intrinsic utility of his own action and the social dissonance associated with the measure of agents choosing this same action.

$$U^i(s) = u^i(s(i)) - d(s(i), \mu_s(s(i))) \tag{3.1}$$

A technical assumption is required: the intrinsic utilities of the players must be distributed in such a way we can determine the measure of players in $[0, 1]$ with an intrinsic utility for a given action that is above a minimum level. This also allows us to measure the set of players who intrinsically prefer one action over another.

Assumption 4. *For each action $c \in C$ and each $a \in \mathbb{R}$, the set $\{i \in [0, 1] : u^i(c) > a\}$ is measurable.*

3.2.1 Dissonance Functions

A dissonance function represents the institution, or exogenous code of conduct, imposed on a society of agents. For each action, the dissonance function defines the psychic cost of each level of agreement with one's action. We might imagine that when a leader makes a statement in support of an action, the dissonance function for that action decreases.

[2] If there are N actions in C, μ_s can be seen as an N-dimensional vector of the measures of agents choosing each action.

Even if the choices of others remained the same, there is a reduction in the legitimacy of peer disapproval and the negative emotion associated with choosing the leader-supported action. Conversely, if more messages from the media and government express a negative stance on a behavior, the dissonance function for the action increases, and peer disapproval is consequently harsher for each level of agreement.[3]

Regardless of the institution, when there is full agreement with others, there is no social dissonance associated with an action. For a fixed dissonance function, as more agents agree with one's own action, the frequency of disapproval (actual or imagined) decreases. It is assumed that there will not be any discrete jumps in the psychic cost of dissonance as a function of the level of agreement. Assumptions 5 and 6 are imposed throughout the paper. The decreasing nature of d and its continuity on $[0, 1]$ guarantees that it attains its maximum for each action when the second argument is zero.

Assumption 5. $d(c, 1) = 0$ *for all* $c \in C$*.*

Assumption 6. *For each* $c \in C$*,* $d(c, \cdot)$ *is non-negative, continuous, and weakly decreasing in the measure of other agents choosing the player's own action.*

We may also impose an additional assumption, that there are no "plateaus" in the graph of the dissonance function until dissonance is reduced to zero. This is both logical and convenient.

Assumption 7. *For all* $c \in C$*,* $d(c, \cdot)$ *is strictly decreasing in the second argument until reaching zero.*

The three assumptions above allow for flexibility in the agreement level at which dissonance is eliminated. The dissonance elimination level may depend on the particular decision or market being studied. For instance, if being in the majority can make an individual impervious to the presence of disagreement, the dissonance elimination level is 0.5.

[3] Diekmann, Przepiorka, and Rauhut (2011) suggests there is a limit to this. A dissonance function can be specified such that minimal dissonance is achieved even without full agreement on one's chosen action.

3.2.2 Nash Equilibrium

Action $s(i) \in C$ is a **best response** to the actions of the other agents, $s \backslash s(i)$, if for all $c \in C$, $u^i(s(i)) - d(s(i), \mu_s(s(i))) \geq u^i(c) - d(c, \mu_s(c))$. A Nash Equilibrium is an action profile, s, where every agent's action is a best response to the actions of the other agents.[4]

In this paper, I focus exclusively on pure strategy equilibria.

Payoff functions are continuous in a player's own action and in the actions of others, and the set $\{i \in [0,1] : U^i(c,s) > U^i(c', s)\}$ is measurable for all $c, c' \in C$ and $s \in S$. These properties and the fact that the dissonance function depends only on the measure, not the names, of the concordant agents allow us to apply Theorem 2 from Schmeidler (1973) to guarantee the existence of an equilibrium in pure strategies.[5]

3.3 Two-Action Games

Suppose the choice set consists of only two actions, A and B. Let $\Delta u^i = u^i(B) - u^i(A)$. Assume the agents are arranged on $[0,1]$ in decreasing order of Δu^i.[6] Throughout the rest of this paper, a player's Δu^i serves as his *type*.

Theorem 9. *(Monotonicity) Suppose s^* is an equilibrium. If j is such that $u^j(s^*(i)) - u^j(c) > u^i(s^*(i)) - u^i(c)$ for $c \neq s^*(i)$ then $s^*(j) = s^*(i)$.*

Proof. Let $\mu_{s^*}(A)$ and $\mu_{s^*}(B)$ be the measure of agents who play A and B in s^*. Suppose that $s^*(i) = B$. This is a best response because s^* is an equilibrium. It must be that $u^i(B) - d(B, \mu_{s^*}(B)) \geq u^i(A) - d(A, \mu_{s^*}(A))$ and hence,

$$u^i(B) - u^i(A) \geq d(B, \mu_{s^*}(B)) - d(A, \mu_{s^*}(A)) \tag{3.2}$$

[4] Others in the literature, including Schmeidler, define Nash equilibrium by allowing a measure-zero set of agents to deviate from the best response, so *almost every* agent is best-responding. Since actions of a measure-zero set of agents do not affect payoffs, there is no substantive different between the two definitions.

[5] For another approach to establishing existence of a pure strategy equilibrium (where *all* agents are best-responding with a pure strategy), see Yang and Qi (2013).

[6] This ordering must be preserved to apply the results of Yang and Qi on monotone equilibria and comparative statics.

If $u^j(B) - u^j(A) > u^i(B) - u^i(A)$ then

$$u^j(B) - u^j(A) > d(B, \mu_{s^*}(B)) - d(A, \mu_{s^*}(A)) \tag{3.3}$$

and only B is a best response for agent j, so $s^*(j) = B$ for all such j.

There is an analogous argument for the case when $s^*(i) = A$ and there is an agent j with $u^j(A) - u^j(B) > u^i(A) - u^i(B)$.

□

From Theorem 9, we know that if Δu^i is strictly decreasing[7] in i, equilibria are strictly monotone: if agent $i \in [0, 1]$ chooses B then so does any agent $j < i$, and if i chooses A then so does any $j > i$. Thus, if Δu^i is strictly decreasing in i, we can describe each equilibrium by a threshold agent, i^*. Agents with intrinsic utility greater than Δu^{i^*} choose action B, and agents with intrinsic utility less than Δu^{i^*} choose A.

3.3.1 Supermodularity

Social dissonance games with a continuum of agents belong to a larger class, nonatomic supermodular games, which are studied by Yang and Qi (2013). The key property of supermodular games is that of increasing differences with respect to actions: the payoff from choosing a particular action increases when a greater measure of other agents also choose that action. Payoffs are also required to be monotone with respect to "type," which in the two-action game refers to the value of Δu^i. Once these properties are verified, the results of Yang and Qi can then be used to establish monotone comparative statics properties.

Theorem 10. *Two action social dissonance games are supermodular.*

Proof. Three properties characterize supermodular games. Firstly, payoffs for each player i are order upper semi-continuous in i's own action for a fixed distribution of other agents'

[7] In cases where there is a positive measure of agents of the same Δu^i, the monotonicity result of Theorem 9 applies, but there may also be asymmetric equilibria. See section 3.4 for discussion.

actions and types, and a fixed Δu^i. Secondly, payoffs must be supermodular in a player's own action when the action is multi-dimensional. These two properties are trivially satisfied since a player's action space is finite and one dimensional.

Thirdly, each player i's payoffs must exhibit increasing differences in own action $s(i)$ and the distribution of other agents' actions and types plus the agent's own type, Δu^i. Social dissonance games defined here are anonymous, in that only the aggregate, $\mu_s(s(i))$, matters and not the types of the other agents who choose the agent's own action, so it suffices to verify that for all i, U^i has increasing differences in other agents' actions and in Δu^i.

Impose the order $B > A$ on each agent's action set[8] . For action profiles, define an analogous, albeit partial, ordering by letting $s \geq s'$ if and only if $s(i) \geq s'(i)$ for almost all $i \in [0,1]$. An agent's payoff function has increasing differences in others' actions if $s \geq s'$ implies that $U^i(B,s) - U^i(A,s) \geq U^i(B,s') - U^i(A,s')$. If this is expanded we can see that the intrinsic utilities on either side of the inequality cancel out, leaving the condition

$$d(A, \mu_s(A)) - d(B, \mu_s(B)) \geq d(A, \mu_{s'}(A)) - d(B, \mu_{s'}(B)) \tag{3.4}$$

Note that when $s \geq s'$, it is also true that $\mu_s(B) \geq \mu_{s'}(B)$ and $\mu_s(A) \leq \mu_{s'}(A)$. Since social dissonance is non-decreasing in level of agreement, $d(B, \mu_s(B)) \leq d(B, \mu_{s'}(B))$ and $d(A, \mu_s(A)) \geq d(A, \mu_{s'}(A))$. These are sufficient to verify (3.4).

Agents' payoffs also exhibit increasing differences in own action and Δu^i. The increasing differences property is evident from observing that, for any fixed profile of others' actions, the incremental benefit of choosing B over A, $U^i(B,s) - U^i(A,s)$, increases when $u^i(B) - u^i(A)$ increases. $\square$

3.3.2 Comparative Statics

The family of two-action social dissonance games, $\{\Gamma(u,d)\}$, is parameterized by the utility functions of the agents and the social dissonance function. Utility functions must be such

[8] This ordering is arbitrary, but defined to be consistent with the definition of $\Delta u^i = u^i(B) - u^i(A)$.

that Δu^i is decreasing in i. The difference in dissonance between the two actions for each aggregate action profile can be summarized by $\Delta d(s) := d(A, \mu_s(A)) - d(B, 1 - \mu_s(A))$. [9]

It is a best response to choose B if and only if $\Delta u^i \geq -\Delta d(s)$ and it is a best response to choose A if and only if $-\Delta u^i \geq \Delta d(s)$. Since a player's best response depends only on the difference in utility between the two actions, Δu^i, and the difference in dissonance, $\Delta d(s)$, it follows that the set of equilibria depends on Δu^i for all $i \in [0,1]$ and $\Delta d(s)$ for all $s \in [0,1]$. We can define a partial order of the parameters by: $(u_1, d_1) \geq_B (u_2, d_2)$ if for all $i \in [0,1]$, $\Delta u_1^i \geq \Delta u_2^i$ and for all $s \in S$, $\Delta d_1(s) \geq \Delta d_2(s)$.

When comparing the equilibrium sets of two games, we compare the smallest and largest equilibria in each set. With respect to the partial order on action profiles defined in Section 3.3.1, the *smallest* equilibrium is $\underline{s}^* \in S$ such that $s^* \geq \underline{s}^*$ for all equilbria s^*. The *largest* equilibrium, $\bar{s}^*$ is defined analogously. Yang and Qi's main result is that the set of monotone equilibria form a complete lattice[10] , and thus, $\underline{s}^*$ and $\bar{s}^*$ are well-defined. For two equilibrium sets $\int_1$ and $\int_2$, $\int_1 \geq \int_2$ if $\inf(\int_1) \geq \inf(\int_2)$ and $\sup(\int_1) \geq \sup(\int_2)$.

If all agents' intrinsic utilities for a particular action increase relative to their intrinsic utilities for the other action, then we would expect that equilibrium choices would shift toward the former action. Likewise, if the social dissonance function for a particular action decreases relative to the social dissonance function for the other action, we would again expect the equilibrium behavior to shift toward the former action. Applicable monotone comparative statics results for finite games appear in Milgrom and Roberts (1990), and for games with a continuum of agents in Acemoglu and Jensen (2010) and Yang and Qi (2013). Below, I apply Yang and Qi's result, but first need to establish some properties of the family of games.

Lemma 11. *Payoff functions U^i have increasing differences in s and (u, d) with respect to the partial order $\geq_B$, holding fixed a player's type and the type-action distribution of other players.*

[9] Note that for any fixed action profile, $\mu_s(B) = 1 - \mu_s(A)$.

[10] If $\mathcal{S}$ is the set of equilibria, then $\inf(\mathcal{S})$ and $\sup(\mathcal{S})$ are included in $\mathcal{S}$.

Proof. Fix a player's type, Δu^i, the types of the other players (which are irrelevant to a player's own payoffs), and the profile of other agents' action. We can show that for any fixed profile of others' actions, s, the incremental benefit of choosing B over A increases when Δd increases.

$$\begin{aligned} U^i(B,s) - U^i(A,s) &= u^i(B) - d(B,\mu_s(B)) - [u^i(A) - d(A,\mu_s(A))] \\ &= u^i(B) - u^i(A) + d(A,\mu_s(A)) - d(B, 1-\mu_s(A)) \\ &= u^i(B) - u^i(A) + \Delta d(s) \end{aligned}$$

□

Proposition 12. *If two games,* $\Gamma(u_1, d_1)$ *and* $\Gamma(u_2, d_2)$ *with equilibrium sets* $\int_1$ *and* $\int_2$ *are such that* $(u_1, d_1) \geq_B (u_2, d_2)$*, then* $\int_1 \geq \int_2$*.*

Proof. As shown in Lemma 11, payoffs have increasing differences in own action $s(i)$ and game parameters (u, d), holding fixed the actions and types of the players. This means that if $\Gamma(u_1, d_1)$ and $\Gamma(u_2, d_2)$ do not have different Δu^i for any i, but $(u_1, d_1) \geq_B (u_2, d_2)$, then the incremental benefit of switching from A to B is higher in Γ.

Consistent with the partial ordering defined above, the distribution of player types is monotone in that when $(u_1, d_1) \geq_B (u_2, d_2)$, the distribution of player types Δu_1 stochastically dominates Δu_2. Since the family of two-action social dissonance games satisfies the increasing differences property and monotonicity of player types, it is a monotone family of supermodular games. Thus, Theorem 2 from Yang and Qi can be applied: both the largest and smallest monotone equilibria rise monotonically with (u, d). Since all equilibria of two-action social dissonance games are monotone, it follows that the largest and smallest equilibria rise monotonically. □

When agents' intrinsic utlities rise with respect to the order $\geq_B$, the largest and smallest equilibria rise monotonically. A change in the equilibrium set may also be due to a change in the dissonance function. A monotone shift in the dissonance function for one action, such as that resulting from a policy change, has a monotone effect on the difference in dissonance, $\Delta d(s)$. Either an increase in $d(A, \mu_s(A))$ for all s or a decrease in $d(B, \mu_s(B))$ for all s would increase $\Delta d(s)$. The following two corollaries of Proposition 12 describe the separate effects of changing either u or d while the other remains the same.

Corollary 1. *If two games, $\Gamma(u_1, d_1)$ and $\Gamma(u_2, d_2)$ with equilibrium sets $\int_1$ and $\int_2$ have the same utility profile ($u_1 = u_2$), and two different dissonance functions d_1 and d_2 such that for all s, $\Delta d_1(s) \geq \Delta d_2(s)$, then $\int_1 \geq \int_2$.*

Proof. Since $d_1 = d_2$, $\Delta d_1(s) = \Delta d_2(s)$ for all s. This fact combined with $\Delta u_1^i \geq \Delta u_2^i$ for all i implies that $(u_1, d_1) \geq_B (u_2, d_2)$. □

Corollary 2. *If two games, $\Gamma(u_1, d_1)$ and $\Gamma(u_2, d_2)$, with equilibrium sets $\int_1$ and $\int_2$ share the same dissonance function ($d_1 = d_2$) and have different intrinsic utility profiles u_1 and u_2 such that for all i, $\Delta u_1^i \geq \Delta u_2^i$, then $\int_1 \geq \int_2$.*

3.4 Games with a Finite Number of Types

This section presents some two-action games where there are a finite number of *types*, or distinct values of Δu^i. First, I consider a game where there are only two types. In two-type games, there may be asymmetric equilibria where agents of the same type choose differently. I also examine games with a third type of agent who is intrinsically indifferent between the two actions.

3.4.1 Games with Two Types

Suppose there are two types of agents: type-b agents with the same $\Delta u^i > 0$ and type-a with the same $\Delta u^i < 0$. Define $\Delta u^b := u^i(B) - u^i(A)$ for type-b's and $\Delta u^a := u^i(A) - u^i(B)$ for type-a's.

Let the measure of type-b agents be some real number $\mu_b \in (0,1)$ and the measure of type-a agents be $\mu_a = 1 - \mu_b$. Arrange the agents on [0,1] by intrinsic utility so type-b agents are in $[0, \mu_b]$ and type-a's are in $(\mu_b, 1]$.

The equilibrium set depends on the distribution of agents in the society, (μ_a, μ_b) and on the strength of agents' preferences relative to social dissonance. Equilibria can be classified as either monolithic, where all agents choose the same action, true-to-type, where each agent chooses his intrinsically preferred action, or asymmetric, where agents of the same type may choose different actions.

Proposition 13. *Sufficient conditions for monolithic and true-to-type equilibria are as follows:*
(1) If $\Delta u^a \leq d(A, 0)$, monolithic-B is an equilibrium.
(2) If $\Delta u^b \leq d(B, 0)$, monolithic-A is an equilibrium.
(3) If $\Delta u^a \geq d(A, \mu_a) - d(B, \mu_b)$ and $\Delta u^b \geq d(B, \mu_b) - d(A, \mu_a)$, true-to-type is an equilibrium.

Proof. (1) If all agents choose B, then B is the only best response for each player of type-b regardless of the dissonance function. The condition $\Delta u^a \leq d(A, 0)$ together with the fact $d(B, 1) = 0$ imply that $u^i(B) - u^i(A) \geq d(B, 1) - d(A, 0)$ for $i \in (\mu_b, 1]$, so B is a best response for type-a agents. The proof of (2) is analogous. (3) If all type-a's play A and all type-b's play B, the measure of agents playing each is μ_a and μ_b. The conditions listed directly imply that A is a best response for type-a's and B is a best response for type-b's □

Certain types of action profiles can be ruled out as candidates for equilibrium.

Corollary 3. *If, in equilibrium, some agents of one type play against-type, then all agents of the other type must play true-to-type.*

Direct proof of this result is provided in the appendix, but it is also a corollary of Theorem 9 on the monotonicity of equilibria. As with the result on monotonicity, the proof rests on the fact that if some type-a's are choosing action B because it offers lower social dissonance, then type-b's are doubly glad to choose B.

Asymmetric equilibria are those where either type-a's or type-b's are split between the two actions, and players of the other type all play true to type. To find equilibria where type-a's behave asymmetrically, compute (I1) the measure of players choosing A and B, $\widetilde{\mu}_a(A)$ and $\widetilde{\mu}_a(B)$,[11] for which a type-a player is indifferent between playing A and playing B. To find equilibria where type-b's behave asymmetrically, compute (I2) $\widetilde{\mu}_b(A)$ and $\widetilde{\mu}_b(B)$ for which type-b's are indifferent between A and B. Under Assumptions 5, 6, and 7, there is a single value of $\widetilde{\mu}_a(A)$ satisfying (I1) and a single value of $\widetilde{\mu}_b(A)$ satisfying (I2).

First, find the value of $\mu(A)$ satisfying (I1), $\widetilde{\mu}_a(A)$. $\widetilde{\mu}_a(A)$ is the value of $\mu(A)$ for which $\Delta u^a = d(A, \mu(A)) - d(B, 1 - \mu(A))$. In the case of social dissonance function

$$d(s(i), \mu_s(s(i)) = max(0, 1 - \frac{\mu_s(s(i)}{\bar{\mu}}) \tag{3.5}$$

this is the value of $\mu(A)$ in [0,1] such that

$$\Delta u^a = max(0, 1 - \frac{\mu(A)}{\bar{\mu}}) - max(0, 1 - \frac{1 - \mu(A)}{\bar{\mu}})$$

The expression on the right side of the equation evaluates as:

$$= \begin{cases} 1 - \dfrac{\mu(A)}{\bar{\mu}} & \text{if } \mu(A) \in [0, min(\bar{\mu}, 1 - \bar{\mu})] \\ \dfrac{1 - 2\mu(A)}{\bar{\mu}} & \text{if } \mu(A) \in [min(\bar{\mu}, 1 - \bar{\mu}), max(\bar{\mu}, 1 - \bar{\mu})]. \\ \dfrac{1 - \mu(A)}{\bar{\mu}} - 1 & \text{if } \mu(A) \in [max(\bar{\mu}, 1 - \bar{\mu}), 1]. \end{cases} \tag{3.6}$$

Second, $\widetilde{\mu}_a(B)$ is the value of $\mu(A)$ for which

$$\Delta u^b = d(B, 1 - \mu(A)) - d(A, \mu(A))$$

[11] Subscript s for the action profile is omitted, since infinitely many profiles have $\mu(A)$ agents playing A and $\mu(B)$ agents playing B.

For the social dissonance function (3.5), this is the value of $\mu(A)$ such that

$$\Delta u^b = max(0, 1 - \frac{1-\mu(A)}{\bar{\mu}}) - max(0, 1 - \frac{\mu(A)}{\bar{\mu}})$$

The expression on the right side of the equation evaluates as:

$$= \begin{cases} \frac{\mu(A)}{\bar{\mu}} - 1 & \text{if } \mu(A) \in [0, min(\bar{\mu}, 1-\bar{\mu})] \\ \frac{2\mu(A)-1}{\bar{\mu}} & \text{if } \mu(A) \in [min(\bar{\mu}, 1-\bar{\mu}), max(\bar{\mu}, 1-\bar{\mu})]. \\ 1 - \frac{1-\mu(A)}{\bar{\mu}} & \text{if } \mu(A) \in [max(\bar{\mu}, 1-\bar{\mu}), 1]. \end{cases} \tag{3.7}$$

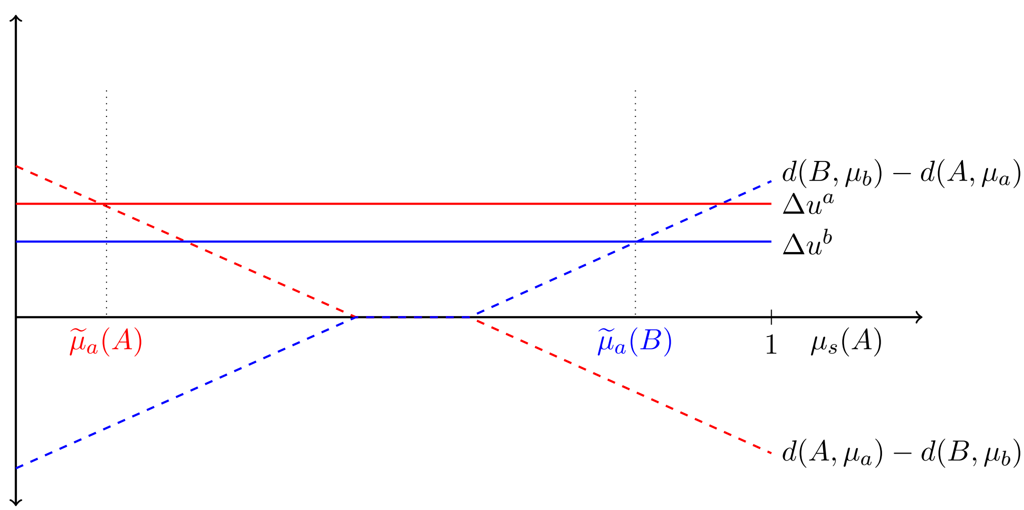

Figure 3.1: Finding asymmetric equilibria

Example 9. Suppose $\Delta u^a = \Delta u^b = 0.25$ and the dissonance function is (3.5) with $\bar{\mu} = 0.75$. To find the equilibria where type-a chooses asymmetrically, find the value of $\mu(A)$ in [0,1] for which

$$0.25 = max(0, 1 - \frac{\mu(A)}{0.75}) - max(0, 1 - \frac{1-\mu(A)}{0.75})$$

Solving this generates a value of $\tilde{\mu}_a(A) = 13/32$. The analogous condition for type-b is

$$0.25 = max(0, 1 - \frac{1-\mu(A)}{0.75}) - max(0, 1 - \frac{\mu(A)}{0.75})$$

and $\tilde{\mu}_b(B) = 19/32$.

We can use these points of indifference to form the interval in which true-to-type equilibrium is possible. In Example 9, the measure of type-a players, μ_a, must be in the interval [13/32, 19/32] in order for a true-to-type equilibrium to be possible. If we find that $\widetilde{\mu}_a(A) = \widetilde{\mu}_a(B)$ then that is the only measure of type-a players that makes a true-to-type equilibrium possible. If $\widetilde{\mu}_a(A) > \widetilde{\mu}_a(B)$ then a true-to-type equilibrium is not possible.

3.4.2 Games with Indifferent Agents

In many matters of public opinion and policy, there are individuals on either side of an issue and those who are removed from it. In addition to type-a and type-b agents, we allow for *type-0* agents who are intrinsically indifferent between A and B; $u^0(A) = u^0(B)$. Let the measures of each type of agent be μ_a, μ_b, and μ_0, with $\mu_a + \mu_b + \mu_0 = 1$. Arrange the agents on [0,1] by intrinsic utility so those of type-b are found in $[0, \mu_b]$ and those of type-a are found in $(1 - \mu_a, 1]$, with type-0 in the middle, $(\mu_b, 1 - \mu_a]$.

Proposition 14. *In an equilibrium with asymmetric choices among type-0 agents, social dissonance for both actions must be equal.*

Proof. If in equilibrium s, both A and B are chosen by type-0 agents, then both are best responses of type-0 agents. Since $u^0(A) = u^0(B)$, this implies that $d(A, \mu_s(A)) = d(B, \mu_s(B))$. □

Under the assumption that $d(c, \cdot)$ is strictly decreasing in the second argument until reaching zero, when indifferent agents act asymmetrically, one of two conclusions is possible. Either there is a singular aggregate $(\mu_s(A), \mu_s(B))$ such that $d(A, \mu_s(A)) = d(B, \mu_s(B))$ or the social dissonance of both actions is zero. In the latter case, action profiles where indifferent agents act asymmetrically are welfare-maximizing.

Consider a society in which there are two ways of life, and in order to feel validated, people require $\bar{\mu}$ of the population to join them, with $\bar{\mu} > \mu_a$ and $\bar{\mu} > \mu_b$. Suppose further that there are $\mu^* \in (0, \mu_0)$ such that $\bar{\mu} \leq min\{\mu_a + \mu^*, \mu_b + \mu_0 - \mu^*\}$. There are two groups who both feel strongly about their preferred way of life, and a group that is indifferent. To maximize the welfare among the population, we would like to encourage selection of an equilibrium where type-a and type-b play true-to-type and type-0 agents split according to a μ^* satisfying the aforementioned criteria.

3.5 Games with a Continuum of Types

Static two-action social dissonance games with a continuum of agent types provide a framework for empirical analysis where there is an incentive to conform. Allowing for a continuous range of Δu^i is more realistic than assuming agents can be categorized into a small number of homogeneous groups.

3.5.1 Assumptions on Functional Form

The best response condition for each agent depends on $\Delta u^i := u^i(B) - u^i(A)$, for each $i \in [0, 1]$, and $\Delta d : [0, 1] \to \mathbb{R}$ which is the difference in social dissonance between choosing B and choosing A, given a proportion of other agents choosing B.

$$\Delta d(\mu_s(B)) = d(B, \mu_s(B)) - d(A, 1 - \mu_s(B)) \tag{3.8}$$

This way of defining Δu and Δd allows us to compare the incremental intrinsic utility of choosing B to the incremental social dissonance of B. An equilibrium can be rationalized by many different configurations of the intrinsic utility and social dissonance functions. To simplify the analysis, we can make some assumptions about the specification for Δu^i and Δd.

Assumption 8. Δu^i *is continuous and strictly decreasing in* i.

Assumption 9. $\Delta u^{i^0} = 0$ *for some* $i^0 \in (0, 1)$.

It seems reasonable that the distribution of intrinsic utilities in a large population is such that for each individual, there is another who is within some small epsilon of intrinsic utility, and that no two individuals have exactly the same intrinsic utilities. It follows from Assumption 8 that Δu^i is bounded. That is, Δu^0 and Δu^1 are within some $[u_{min}, u_{max}]$. Assumption 9 focuses us on the interesting cases where there is a positive measure of people who have a clear intrinsic preference for A and of those who have a clear intrinsic preference for B.

One of the simplest specifications for d that satisfies Assumptions 5, 6, and 7 has dissonance decreasing linearly in measure of agreement from its maximum, until reaching zero at some measure of agreement that may be less than 1.

Assumption 10. *The dissonance function, d, takes the following form:*

$$d(s^i, \mu_s(s^i)) = M_{s^i} \cdot max(0, 1 - \frac{\mu_s(s^i)}{e_{s^i}}). \tag{3.9}$$

For each action, there are two parameters. One is the maximum dissonance, $M_{s^i} > 0$, which is experienced when no other agents agree with one's own action. The second parameter is the dissonance elimination level, $e_{s^i} \in (0, 1)$, or the minimum measure of agreement from other agents needed for a zero level of dissonance. The dissonance functions for each of the two actions can be combined into a single function Δd, written as a function of the measure of agents choosing B, $\mu_s(B)$. There are two main cases to deal with, based on the values of e_A and e_B.

Social dissonance is zero for both actions when $\mu_s(B) \geq e_B$ and $\mu_s(A) \geq e_A$. These conditions together require that $e_B \leq 1 - e_A$. If this inequality is strict, there is a portion of the graph of Δd that is flat and the function can be written as:

$$\Delta d(\mu_s(B)) = \begin{cases} M_B \cdot (1 - \frac{\mu_s(B)}{e_B}) & \text{if } \mu_s(B) \leq e_B \\ 0 & \text{if } e_B \leq \mu_s(B) \leq 1 - e_A \\ -M_A \cdot (1 - \frac{1-\mu_s(B)}{e_A}) & \text{if } \mu_s(B) \geq 1 - e_A \end{cases} \tag{3.10}$$

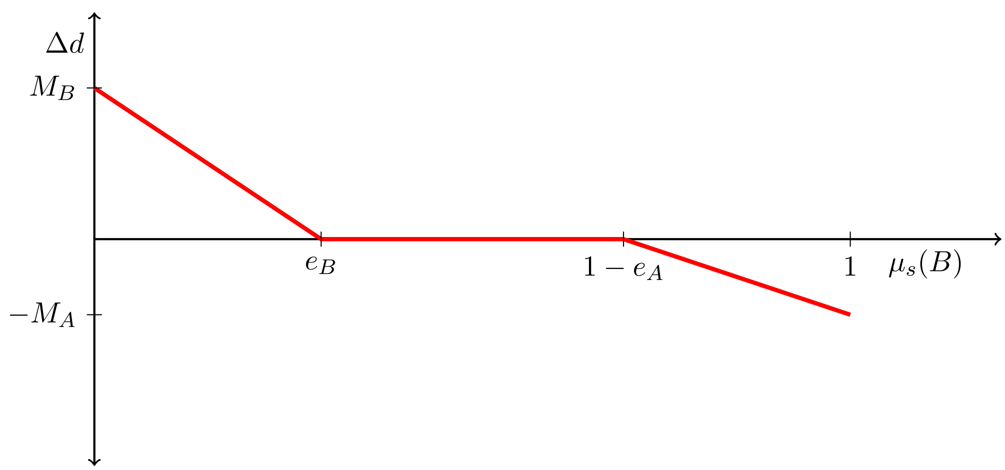

Figure 3.2: An example of Δd following the form of 3.10

In the case that $e_B \geq 1 - e_A$, Δd is strictly decreasing in $\mu_s(B)$ over $[0, 1]$, and there is a single value of $\mu_s(B)$ for which $\Delta d = 0$. Δd can be written as:

$$\Delta d(\mu_s(B)) = \begin{cases} M_B \cdot (1 - \frac{\mu_s(B)}{e_B}) & \text{if } \mu_s(B) \leq 1 - e_A \\ M_B \cdot (1 - \frac{\mu_s(B)}{e_B}) - M_A \cdot (1 - \frac{1-\mu_s(B)}{e_A}) & \text{if } 1 - e_A \leq \mu_s(B) \leq e_B \\ -M_A \cdot (1 - \frac{1-\mu_s(B)}{e_A}) & \text{if } \mu_s(B) \geq e_B \end{cases} \tag{3.11}$$

Observe that in both of these cases, the function Δd takes on a nonzero value for at least one value of $\mu_s(B)$. This, in combination with Assumptions (8) and (9), guarantees there is some positive measure of agents for which choosing their intrinsically preferred

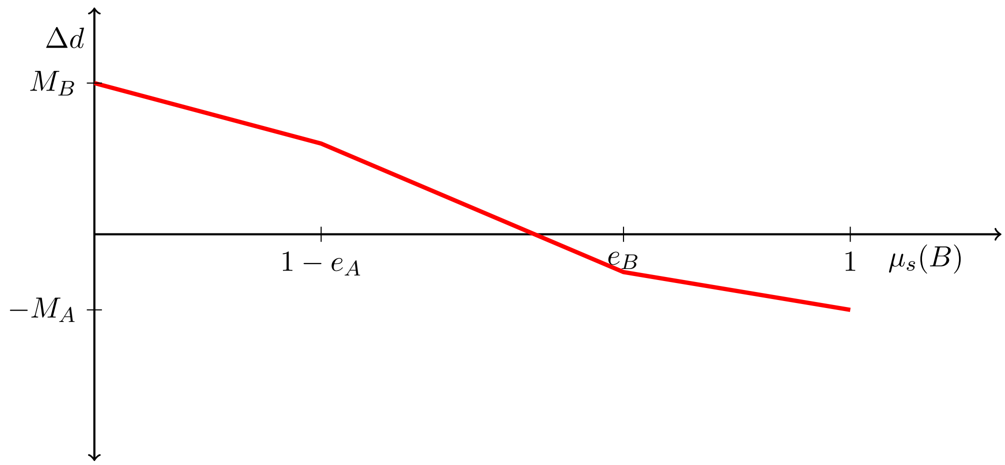

Figure 3.3: An example of Δd following the form of 3.11

action is not a dominant strategy. Hence, these agents must consider what others are planning to do.

We might argue that case 2 is much more common in reality; if we assume that being in the minority causes a positive level of social dissonance, this is $e_A, e_B \geq 1/2$.

3.5.2 Equilibrium

Equilibria are monotone under Assumptions 5, 6, and 7, which is to say that the players who choose B are those with the strongest intrinsic preference for it. In equilibrium, players on $[0, 1]$ to the left of some threshold choose B, and those to the right of the threshold choose A.

A **monolithic** action profile is a monotone action profile in which all agents choose the same action. These can be studied separately, and conditions for monolithic equilibria are provided in section 3.4 and a previous essay. Here, we focus mainly on identifying non-monolithic equilibria, which are action profiles s^* with a positive measure of agents choosing either action, such that for (almost) all $i \in [0, \mu_{s^*}(B)]$,

$$\Delta u^i \geq \Delta d(\mu_{s^*}(B)) \tag{3.12}$$

and for (almost) all $i \in [1 - \mu_{s^*}(B), 1]$,

$$\Delta u^i \leq \Delta d(\mu_{s^*}(B)) \tag{3.13}$$

Together these imply that for each equilibrium s^*, there is a threshold agent, i^* such that $\Delta u^{i^*} = \Delta d(\mu_{s^*}(B))$ and since $\mu_{s^*}(B) = i^*$, this condition is

$$\Delta u^{i^*} = \Delta d(i^*) \tag{3.14}$$

Depending on the functional form, there may be multiple non-monolithic equilibria.

3.5.3 Basic Linear Specification

The specifications in the previous sections impose boundedness and linearity on Δu^i and piecewise linearity on Δd. Here, I specify a model which allows for only the most basic features.

Intrinsic utility, through the mapping $\Delta u^i : [0,1] \to \mathbb{R}$ should allow for a positive measure of two types of agents: those who intrinsically prefer A and those who intrinsically prefer B. One way to do this is to allow for only two values of Δu^i: a positive value and a negative one. Another way to do this, consistent with Assumption 8, is to let Δu^i be linear in i, as in Dixit (2003).

$$\Delta u^i = \beta - \alpha i \tag{3.15}$$

with $\alpha > \beta > 0$. The measure of agents who prefer B is $i^0 = \beta/\alpha$ and the measure who prefer A is $1 - \beta/\alpha$. [12] If the assumption $e_A = e_B = 1$ is imposed on Δd, equation 3.11 reduces to

$$\Delta d(\mu_s(B)) = M_B - \mu_s(B)(M_A + M_B) \tag{3.16}$$

[12] This can be further simplied to $\Delta u^i = \beta - i$ if one wishes to reduce the number of parameters.

This is again similar to the specification in Dixit (2003) and can be written as [13]

$$\Delta d(\mu_s(B)) = \sigma - \tau \cdot \mu_s(B) \tag{3.17}$$

Using the specification given in (3.15) and (3.17) in threshold condition (3.14), the threshold agent in a non-monolithic equilibrium is

$$i^* = \frac{\sigma - \beta}{\tau - \alpha} \tag{3.18}$$

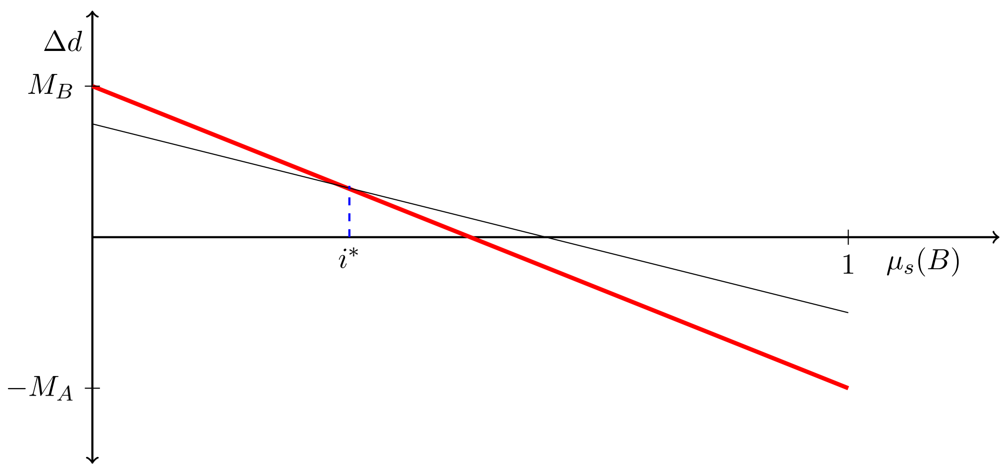

Figure 3.4: Finding the threshold agent in the basic linear model

3.5.4 General Linear Specification

If we remove the assumption $e_A = e_B = 1$, and allow for these parameters to be any values in $(0, 1)$, Δd will follow either (3.10) or (3.11), but the logic for finding the set of non-monotone equilibria is the same.

There may be a portion of the graph of Δd that is flat, but this does not complicate things much. The set of potential equilibria is limited by the fact that when dissonance

[13] An even simpler specification would be $\Delta d(\mu_s(B)) = \sigma - \mu_s(B)$. For this, we would require $0 < \sigma < 1$ so there is a minimum measure of agents choosing A, between 0 and 1, for which the dissonance of B is at least as great as the dissonance of A.

is zero for all agents, and Assumptions (8) and (9) hold, there is a single agent who is indifferent between the two actions. If $\Delta d = 0$ in equilibrium, we have $i^* = i^0$[14], with $s^*(i) = B$ for $i < i^*$ and $s^*(i) = A$ for $i > i^*$.

Otherwise, the computation of the threshold agent is similar to that in the case where $e_A = e_B = 1$; we just have to consider more cases where Δu^i may intersect Δd.

In case 1, monolithic-A is not an equilibrium, since there are agents for whom B is a dominant strategy. In fact, monolithic-B is the only equilibrium in this type of situation.

In case 2, monolithic-A and monolithic-B are both equilibria. To see why monolithic-A is an equilibrium, consider the decision of the agent $i = 0$ whose incremental benefit from choosing B is outweighed by the incremental dissonance of going it alone. If the agent of strongest intrinsic preference for B is not willing to deviate from A, then none of the other agents will. A true-to-type action profile is also an equilibrium; all agents, if choosing true-to-type, face zero dissonance and have no reason not to choose according to intrinsic preference. Notice there are two other points where Δu crosses Δd, and for $i^* = \mu_s(B)$, satisfy equilibrium conditions (3.12) and (3.13).

In case 3, both monolithic-A and monolithic-B are ruled out as equilibria, since for each action there is an agent who would be willing to choose it alone. Only true-to-type is an equilibrium.

[14] If the threshold agent is i^0 the equilibrium is true-to-type.

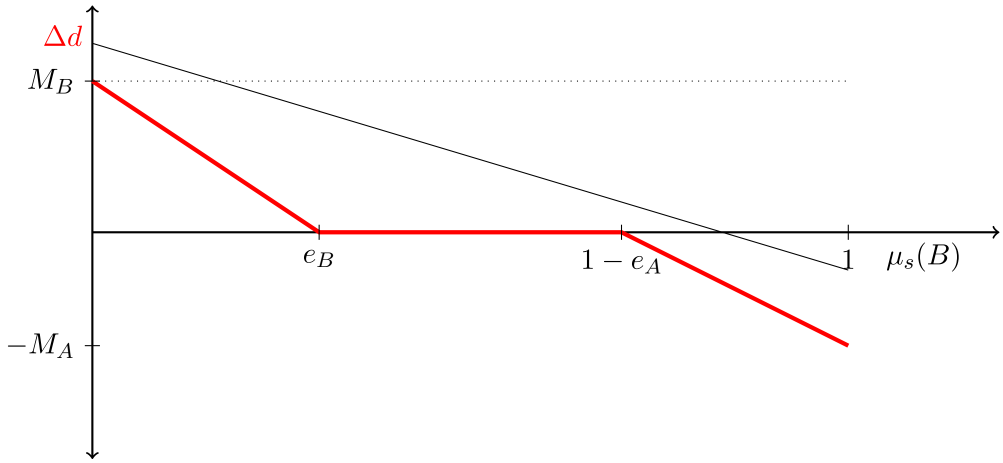

Figure 3.5: Case 1: Single monolithic equilibrium

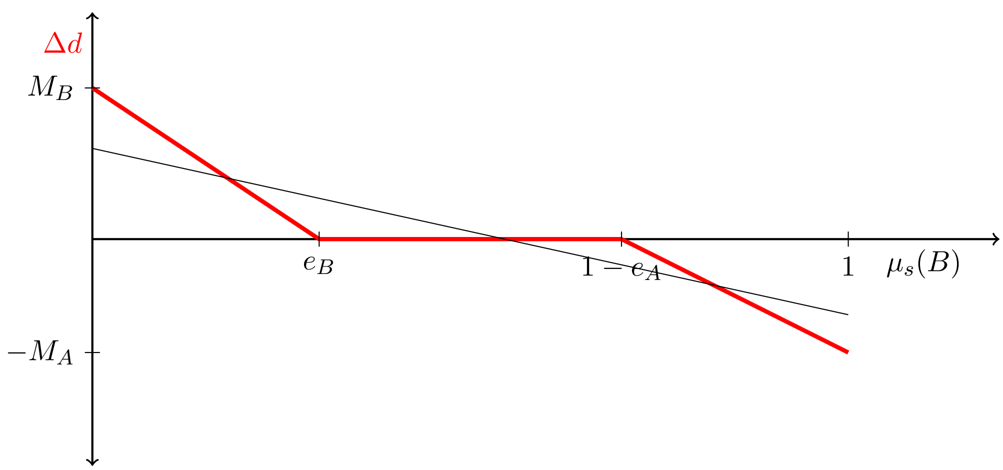

Figure 3.6: Case 2: True-to-type and both monolithic equilibria

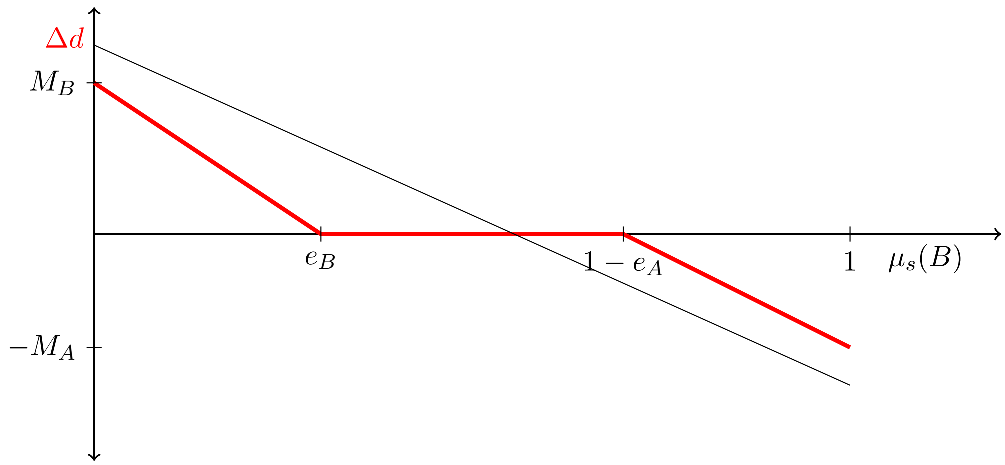

Figure 3.7: Case 3: Single true-to-type equilibrium

The above cases are not an exhaustive taxonomy of possible situations, but should provide an idea of how to determine the equilibrium set from a given $(\Delta u, \Delta d)$ pair.

3.6 Social Dynamics

Trends for behavior (and opinion) throughout history often follow a pattern where one behavior is dominant, there is a shift to another behavior, and that second behavior remains dominant until the next shift. Before the shift, the previously dominant behavior is chosen by a relatively steady proportion of the population, and after the shift, the proportion of the population choosing the newly dominant behavior approaches its own steady state.

The static model of social dissonance games provides three explanations for a shift in observed behavior. Two of these are discussed in section 3.3.2 on comparative statics. Intrinsic utilities may change and social dissonance functions may change. Another possibility is that a different equilibrium is selected from a set of multiple equilibria. The literature on global games, including Frankel, Morris and Pauzner (2003) addresses this mechanism in a game with noisy information. In games of complete information, there is little guidance about how one equilibrium is selected, though Bisin, Moro and Topa (2011) provide a framework for estimating parameters in network games with multiple equilibria.

3.6.1 Challenges in Identification

If we assume a single observation represents an equilibrium of a social dissonance game, there is an infinite set of possible intrinsic utility and dissonance functions that can explain the observation. Too much flexibility can lead to "overfitting;" nonlinearities and plateaus in the dissonance function, for instance, need to be justified if the model is to have predictive power. Even assuming that a series of observations $i_1^*, i_2^*, ..., i_T^*$ fits the basic linear specification of section 3.5.3, the functions Δu and Δd are defined in terms of four parameters: α, β, σ and τ, each of which can change over time. In order to identify these parameters, some additional assumptions may be needed about which parameters change from one observation to the next.

As an example, consider the rise of cohabitation of unmarried couples in the United States. State by state, divorce laws shifted to a no-fault system where partners may agree to end a marriage based on "irreconcilable differences" rather than a fault-based system where divorce can only be triggered by infidelity, abuse, or protracted separation (Vlosky and Monroe, 2002). No-fault divorce laws in the United States were enacted as early as 1969, but the vast majority went into effect between 1970 and 1979. It can be argued that no-fault divorce lowers the sanctity of marriage, hence lowering the social dissonance of cohabitation as a marriage alternative for couples in committed relationships.

Cohabitation increased following the shift to no-fault divorce, but the increase cannot be pinned on this clearly identifiable "shift in institutions." In most states, cohabitation was also increasing leading up to the passage of the law, so perhaps the "institution" was already shifting in advance of the anticipated legal change. Additionally, there is a learning explanation - in the spirit of Fogli and Veldkamp (2011) - as more couples cohabit, the costs and benefits become more apparent to others who previously followed the status quo because of ambiguity about cohabitation-related outcomes.

In this example, three separate explanations have already been proposed: the institution shifted when the law was changed, the institution was shifting before the law was changed, and there was a shift in intrinsic utilities. Surely, combinations of these as well as other explanations are possible. The monotone comparative statics results of section (3.3.2) support many hypotheses, and it is even possible that either Δu or Δd moves in a counterintuitive direction, yet is offset by a movement in the other set of parameters. Ahead, I consider the consequences of highly restrictive assumptions about functional form and changes in parameters.

3.6.2 Shifting Institutions and Drifting Intrinsic Utilities

Suppose that the basic linear specification of (3.15) and (3.17) is a reasonable representation of the distribution of intrinsic utilities and social dissonance. The equilibrium proportion of agents choosing B (3.18) is a function of the parameters α, β, σ, and τ. Suppose further that we assume intrinsic utilities drift over time, by some learning process, whether or not the institution shifts. Holding dissonance parameters τ and σ constant, the effect of a change in α from one observation to another (holding β fixed) is

$$i_2^* = \frac{\tau - \alpha_1}{\tau - \alpha_2} i_1^* \tag{3.19}$$

If intrinsic utilities continue to drift to α_3 and the institution lowers social dissonance for B so $\tau_3 < \tau_2$, the effect on equilibrium is:

$$i_3^* = \frac{\tau_3 - \sigma + \beta}{\tau_2 - \sigma + \beta} \cdot \frac{\tau_2 - \alpha_2}{\tau_3 - \alpha_3} i_2^* \tag{3.20}$$

There is, of course, the question of whether a shift in intrinsic utilities will apply more pressure to policymakers and other leaders to shift an institution. In many cases, the shift in i^* before the shift in institution will be in the same direction as the shift in i^* after the institution changes. In other cases, the institution shifts in the opposite direction as an observed shift in i^*. For example, countermeasures to increase social dissonance may be enacted when strategic defaults become too common or if public smoking becomes widespread.

3.7 Dynamic Games

Consistency is a valued trait. Politicians often suffer from a damaged reputation for changing their stances on issues. People can also be subject to psychic costs associated with changes in opinion or behavior, representing an internal inconsistency (Festinger, 1962) between one's past and present selves. These *switching costs* can prevent a good idea from being adopted, or cause inertia when intrinsic utilities shift to favor one action over another. However, if agents anticipate that an institution will shift, they may shade their expressed opinions in the direction of the shift to reduce social dissonance.

Dissonance in period t depends only on the action profile for period t. The switching cost, $c(s_t^i, s_{t-1}^i)$, is the cost of taking action s_t^i given that action s_{t-1}^i was chosen in the previous period. Assume $c(s_t^i, s_{t-1}^i) > 0$ if $s_t^i \neq s_{t-1}^i$ and $c(s_t^i, s_{t-1}^i) = 0$ if $s_t^i = s_{t-1}^i$. The utility in period t for agent i, given current action profile s_t and the previous period's action profile s_{t-1}, is

$$U_t^i(s_t, s_{t-1}) = u^i(s_t^i) - d(s_t^i, \mu_{s_t}(s_t^i)) - c(s_t^i, s_{t-1}^i) \tag{3.21}$$

Thus, the payoffs in a dynamic social dissonance game with switching costs, assuming no discounting and no changes in the intrinsic utilities or dissonance function, are

$$U^i(\{s_t\}_{t=0}^T) = \sum_{t=0}^{T}[u^i(s_t^i) - d(s_t^i, \mu_{s_t}(s_t^i))] - \sum_{t=1}^{T} c(s_t^i, s_{t-1}^i) \tag{3.22}$$

3.7.1 Two-Period Model

In a two-period model, the payoffs to agent i are

$$u^i(s_1^i) + u^i(s_2^i) - d(s_1^i, \mu_{s_1}(s_1^i)) - d(s_2^i, \mu_{s_2}(s_2^i)) - c(s_2^i, s_1^i) \tag{3.23}$$

where s_1 and s_2 are the action profiles of the agents in periods 1 and 2.

There may be equilibria where all agents choose the same action in both periods and equilibria where some agents switch. Due to positive switching costs, if a first-period action profile is an equilibrium in the one-period game, and the other agents do not switch their actions in the second period, then an agent's best response is to continue with the same action. However, if there is an expectation that the actions of others will change and there is a sufficiently lowered dissonance from making the same change, then agents may be willing to bear the switching cost.

To simplify the discussion, consider a society with two types of agents, as in section 3.4.1, and a switching cost, $c > 0$, when $s_1^i \neq s_2^i$. When agents do not switch, the payoffs are $2u^i(s_1^i) - 2d(s_1^i, \mu_{s_1}(s_1^i))$. As in the static model, some agents face a conflict between the motives to maximize intrinsic utility and to reduce social dissonance. The equilibrium set may include monolithic-A, monolithic-B, true-to-type, or the same asymmetric equilibrium played twice. Conditions for these are found in section 3.4.1.

There are many candidates for two-period equilibria. Focusing only on the symmetric action profiles, there are nine possible combinations of monolithic-A, monolithic-B, and true-to-type. If a two-period action profile is an equilibrium, there can be no deviations from that profile that would improve the payoffs of either type of agent. For example, if (monolithic-A, true-to-type) is an equilibrium, it must be that

$$\begin{aligned} 2\Delta u^a - d(A, \mu(A)) &\geq \Delta u^a - d(B, 1 - \mu(A)) - c \\ 2\Delta u^a - d(A, \mu(A)) &\geq -d(B, 0) - d(B, 1 - \mu(A)) \end{aligned} \tag{3.24}$$

and

$$\begin{aligned} \Delta u^b - d(B, 1 - \mu(A)) - c &\geq -d(A, \mu(A)) \\ \Delta u^b - d(B, 1 - \mu(A)) - c &\geq 2\Delta u^b - d(B, 0) - d(B, 1 - \mu(A)) \end{aligned} \tag{3.25}$$

The first pair of conditions compares a type-a player's payoff from choosing A in both periods to the payoffs from choosing (A,B) and (B,B). A deviation to (B,A) is dominated by (A,A) because it has a lower intrinsic utility, higher switching cost, and higher dissonance.

The last two conditions compare a type-b player's payoff from choosing A in the first period and B in the second to (A,A) and (B,B). (B,A) is dominated by (A,B).

Example 10. Suppose there is a society which is half type-a and half type-b. We have $u^a(A) = u^b(B) = P$, where $0 < P < 1$, $u^a(B) = u^b(A) = 0$, and $d(s_t^i, \mu_{s_t}(s_t^i)) = max(0, 1 - \frac{\mu_{s_t}(s_t^i)}{0.5})$ for $t = 1, 2$. In the stage game (without switching cost), the equilibria are true-to-type, monolithic-A, and monolithic-B. There are also asymmetric equilibria where one type is evenly split between A and B. When there is no switching cost, we can see that the equilibria in the two-period game are formed by selecting pairs from the set of the one-period game's equilibrium set.

However, if switching cost, C, is strictly positive, then some of the pairs of one-period equilibria may be excluded from the set of two-period equilibria. For a two-period strategy profile of (monolithic-A, monolithic-B) or (monolithic-B, monolithic-A) the payoff to a player of either type is $P - C$. If a player of either type deviates unilaterally to choose his intrinsically preferred action in both periods, the switching cost is avoided, but a social dissonance cost of -1 is incurred for a total payoff of $2P - 1$. Thus (monolithic-A, monolithic-B) and (monolithic-B, monolithic-A) are equilibria if and only if $P + C \leq 1$. A similar analysis reveals that profiles involving one period of monolithic choices and one period of true-to-type choices are equilibria if $P + C \leq 1$ and $P \geq C$.

3.8 Conclusion

This essay extends the model of finite-agent social dissonance games to games with a continuum of agents. This represents a more natural way of handling aggregate-level data. Supermodularity, monotonicity of equilibria and monotone comparative statics properties carry over to the continuum-of-agents model. Due to the individual agent's lack of influence, asymmetric equilibria are possible among non-indifferent agents.

There are many modeling choices that need to made in advance of using the social dissonance paradigm to measure or predict the effect of a policy. Evidence from psychology can help indicate the general shape of social dissonance functions, including degree of concavity and the point of dissonance elimination. Knowledge of mental and social processes can also aid in understanding when a dynamic game, with its psychic cost of switching one's action or opinion, is appropriate.

Identification of parameters may require additional assumptions about the baseline rate of learning in a population and how it affects trends in behavior, as well as awareness of how changes in institutions are anticipated. There is an additional element, unexplored here, of how opinions and behavior can push institutions to change, which in turn influences opinions and behavior. In my model, I have treated social dissonance and intinsic utilities as two separate entities, but I hope that future work will successfully integrate the two, while recognizing the flexible boundary between the individual and society.

Chapter 4

Impulse and Temptation

4.1 Introduction

People feel impulses to consume unhealthy things. These impulses are often brought on by sensory cues, but can also arise spontaneously from a train of thoughts or a physiological process. If I mention ice cream, you might get an impulse to consume ice cream. You resolve this impulse by weighing the imagined pleasure you would gain from consuming ice cream against the costs: transportation to the nearest ice cream retailer, some money, and some negative consequences for your health. If you are able to overcome the impulse to consume ice cream, perhaps you are better able to resist it in the future.[1]

Your method of dealing with the impulse is important to future consumption. If you go to the ice cream parlor and purchase one ice cream cone, the impulse is sated for now with no further temptation to consume. If, however, you go to the store and buy a carton of ice cream, this will both tempt you to consume more in the present period, and provide a low cost way of fulfilling future impulses. Even if you are aware that you will face further

[1] Another view is the opposite, that we have a stock of "cognitive control" that gets depleted each time we override an impulse.

temptation to consume, the low per-unit cost of buying a carton may make it seem a better decision than buying a single serving of ice cream.

Cigarette smokers face impulses but do not often have a choice about the quantity of cigarettes they purchase. In the U.S., a pack contains twenty cigarettes, and cartons contain ten packs. Smaller packs are prohibited, and single cigarettes, or "loosies," are not made available by cigarette manufacturers. The World Health Organization (WHO) advises governments to ban the sale of loosies, and many follow this recommendation. One of the reasons cited for this stance against single cigarettes is that their availability encourages consumption by minors and young adults, who are thought to be more susceptible to habit formation than older adults (Ling and Glantz, 2002).

For consumers who are already addicted to cigarettes, or those who are occasional smokers that may become addicted, single cigarettes represent a way to fulfill impulses while limiting the temptation to consume further. Formerly-quit smokers may describe how, overcome by an urge for "just one puff," the purchase of a full pack leads to a complete relapse into the old habit. One might argue that by forcing smokers to choose between no cigarettes and a full pack, the high price of the pack and the knowledge of one's self-control problem would induce the would-be smoker to forgo smoking altogether. This is optimistic, though, knowing the overwhelming strength of the compulsion when it is triggered. Single cigarettes allow for fulfillment of intense impulses without the temptation of the remaining pack.

Many smokers buy packs instead of cartons as a self-control measure. Smokers also report using loosies for self-control purposes (Thrasher et al., 2009), even though the per-cigarette cost is nearly twice what is paid when buying a pack.[2] There is evidence that larger portions of food in restaurants induce diners to eat more, even though taking home leftovers is an option (Rolls et al. 2002). Wansink et al. (2005) offer the explanation that visual cues of portion sizes suggest larger consumption norms. It is possible that the pack

[2] Buyers of single cigarettes also name convenience as a reason for their choices, so perhaps transaction costs are a significant factor preventing impulsive purchases.

size of twenty cigarettes suggests a consumption norm that smokers are inclined to follow. Nearly 30% of smokers report smoking exactly twenty cigarettes per day. Even with some measurement error, it appears that smokers in the U.S. anchor their daily consumption to the size of the pack. Figure 4.1 shows the distribution of number of cigarettes smoked by current smokers[3] with a bin size of 5.

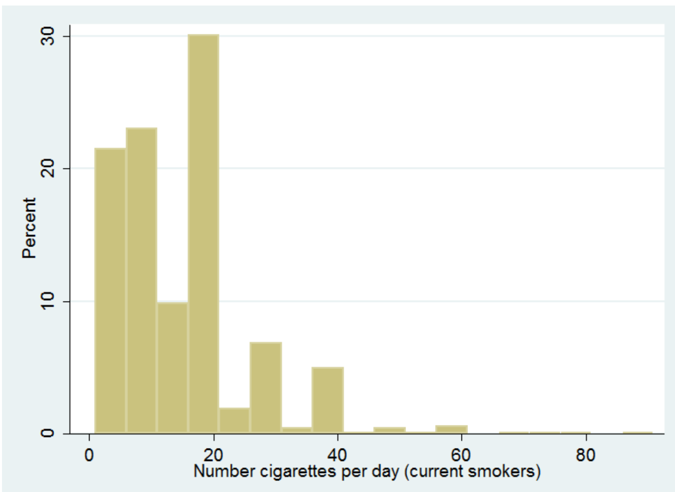

Figure 4.1: Self-reported number of cigarettes smoked per day for current smokers.

Cigarette manufacturers, like producers of other consumables, decide how to bundle their products to maximize revenue in the present and future periods. The choice of twenty cigarettes per pack is perhaps practical – it reduces packaging and transaction costs – and perhaps strategic – knowing that consumers can easily become tempted and addicted, a larger pack provides impetus for further consumption by the otherwise one-time smoker. This can increase future revenue far beyond the revenue lost from would-be infrequent smokers dissuaded by the high price of a pack.

Producer choice is not the only factor determining how cigarettes are sold to consumers.

[3] Data retrieved from Minnesota Population Center, Integrated Health Interview Series. Survey was conducted in 1997. The distribution of daily smoking in 2001 is similar, even in the face of a greater than 50% increase in the per-pack price.

Policy is also relevant. By using a model of decision-making with temptation inspired by Gul and Pesendorfer (2001, 2007), I investigate the question of how the availability of single cigarettes affects the incidence and intensity of smoking. That is, how does the policy banning single cigarettes separately influence the number of smokers and the number of cigarettes each smoker consumes? In a one-period model in which addiction is irrelevant and consumers are aware of the temptations of a larger pack, a policy which removes the small or single-serving pack from consideration can only be welfare-decreasing. Since the one-period model is unrealistic, I investigate separately the effects of addiction and naivete about temptation, each of which may justify the policy.

Naivete about temptation does not lead to mistaken purchase of the small pack, and removal of the small pack from the choice set does not reduce the distortionary effect of naivete on the valuation of the large pack relative to its closest alternative. Therefore, it does not appear that naivete about temptation is in itself a sufficient justification for prohibiting sale of the small pack.

The effect of prohibiting the small pack is less clear in the two-period setting. Some consumers who would otherwise purchase a small pack will instead purchase a large pack and abstain from consuming further, which is inefficient, but as in the one-period model, there will be consumers who avoid purchase altogether. While it is not formally addressed here, it is possible that mistaken beliefs about addiction can justify a government policy restricting the sale of the small pack. I have set up the framework of accounting for this question, but it is up to the empirical researcher to gather data on actual and forecasted addiction.

4.2 Literature

Pollak (1970) is one of the earliest authors to point out and model the fact that short-run consumption affects long-run consumption through habit formation. His utility functions have a "psychological need" parameter which is a function of consumption in each of the past periods; I model psychological needs by tracking the distribution of impulses in the present period as a Markov process.

Becker and Murphy (1988) present a model of rational addiction where consumers have complete information about how present consumption will affect the utility of future consumption. My model focuses on the effects of menu[4] choice and consumption on impulses, and allows agents to be misinformed or uncertain about the effects of their consumption choices on future impulses.

Bernheim and Rangel (2004) construct a model in which decision makers operate in either a cold mode, in which choices are made by cognitive control to maximize utility, or a hot mode, where a strong impulse overwhelms cognitive control to induce use of an addictive substance. The hot mode is stochastically triggered, with a probability that depends on the activity chosen while in a cold mode. Decision makers may choose an *environment* that is absent of consumption cues; my model allows consumers to block themselves from having a ready supply of a tempting substance.

Gul and Pesendorfer (2001) introduce *temptation preferences with self-control* where an agent incurs a cost from resisting a tempting second-best alternative that the ex-ante self found inferior. Gul and Pesendorfer (2004) extends the model of self-control and temptation to dynamic infinite-horizon decision problems, and Gul and Pesendorfer (2007) uses this model to incorporate addiction. In the latter paper, the effects of government price policy and policy on maximal consumption are analyzed separately. The main question I address in this essay is not about the welfare effects of limiting maximal consumption, but rather, setting a minimal purchase quantity.

4.3 Theoretical Framework

The framework of Gul and Pesendorfer (2001) provides inspiration for the model of consumer temptation that follows. In their model, larger menus of different goods require a

---

[4] Going forward, I use menu instead of pack to be consistent with existing literature on temptation.

higher cost of self-control to resist a tempting alternative. Since there is only one good in my model, the cost of temptation is reflected in the increased likelihood of a strong impulse when a larger menu has been selected. In Gul and Pesendorfer (2007), the study of temptation is extended to a multi-period setting which allows for consumption-dependent addiction that increases the intensity of future impulses.

Before describing the timing of consumer impulses and decisions, I will say a word about menus. A *menu* is used in this paper to describe a set of possible consumption plans of a single good in a single time period. Once a menu has been chosen by a consumer at the beginning of a time period, the choices available to the consumer are fixed. There is no opportunity to purchase more of the good, or to change to a different menu. The consumer may choose to consume less than the maximum quantity available and discard the remaining items.

There are three menus available to the consumer: the empty menu ($\emptyset$), a small menu (S), and a large menu (L). Within each time period, there are two opportunities to consume a single unit of the good. These two consumption choices are c^1 and $c^2 \in \{0,1\}$, thus $\emptyset = \{(c^1, c^2) = (0,0)\}$, $S = \{(c^1, c^2) : c^1 + c^2 \in \{0,1\}\}$, and $L = \{(c^1, c^2) : c^1 + c^2 \in \{0,1,2\}\}$. The model could conceivably be extended to accommodate more opportunities for consumption within a single time period and hence, more choices of menus.[5]

4.3.1 Consumers in the One-Period Model

The timing of a consumer's decision-making within each period is as follows: first, the consumer experiences a stochastic impulse to consume one unit of the good. Individuals may differ in the intensity of impulses they experience. An impulse, I, which represents how much the consumer is willing to pay in pecuniary and health costs to remove the urge to consume, is drawn from some distribution Φ with support in $[0, \infty)$.

After the stochastic impulse is realized, the consumer enters the purchase phase, where

[5] For example, if one wishes to directly study the policy on cigarette packages, the largest menu could contain twenty and any smaller menu could also be available.

a menu, m is chosen from a menu of menus, $\mathcal{M} = \{\emptyset, S, L\}$. $\mathcal{M}$ and the prices, p_m, for each $m \in \mathcal{M}$ are determined exogenously by manufacturers, retailers, and regulations. The consumer may also choose the empty menu and forgo purchase. Immediately after purchasing a non-empty menu, the consumer fulfills the initial impulse by consuming one unit of the good.[6]

Following the purchase and initial consumption decisions, the consumer enters a temptation phase, where a second, usually less intense impulse, $I' \in [0, \infty)$ is drawn from another distribution, $\Phi'(m)$ that depends on the remaining quantity of goods available on the menu, and hence, the menu chosen. This temptation is potentially exacerbated in a "use it or lose it" scenario.[7] For example, if a consumer has an impulse for a hamburger, and the vendor offers a "buy one, get one free" special, the weak temptation for the second hamburger and/or aversion to discarding it, can lead to consumption in excess of what the consumer needed to fulfill the initial impulse. A *tempting good* is defined as one where the consumer's impulse to consume it becomes stronger once it is in his possession. The reasons for this could be many; disposal aversion (Bolton and Alba, 2012) is but one possible motivation. The key assumption of my model is that temptation-phase impulses are more likely to be intense when a large menu was chosen.

Assumption 11. *For two menus m and $\hat{m}$ with $m \subset \hat{m}$, $\Phi'(\hat{m})$ first-order stochastically dominates $\Phi'(m)$.*

After the temptation phase, ex post utility is tallied. For unhealthy goods that represent a small portion of the budget, we can assume utility is linear in money. Hence, I proceed by ignoring the utility of normal goods.[8] Unfulfilled impulses are treated as costs. In each

[6] This assumption may be relaxed, but is based on a premise that the initial impulse is most intense, so the main goal of purchase is to fulfill that initial impulse, rather than defer consumption. Assumption 15 frames this for a discrete impulse distribution.

[7] In this paper, I do not allow for saving. But even in the case when saving is possible, the availability of the saved menu items could increase the strength of the initial impulse in future periods and lower the marginal cost of fulfilling those impulses.

[8] The price of the menus enters the decision-making process through the consumer's maximization of surplus.

period the consumer's utility is separable into a disutility of unfulfilled impulses I and I', and a health cost H. Utility can be written as

$$U(c^1, c^2; I, I') = (1 - c^1)(-I) + (1 - c^2)(-I') - H(c^1 + c^2) \tag{4.1}$$

where the realization of I' is drawn from a distribution that depends on m.

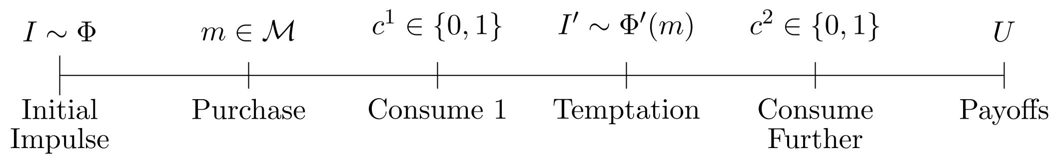

Figure 4.2: Timing of impulses and decisions in the one-period model

The one-period value of a menu is determined by the consumption choices from it that maximize expected utility, with respect to the initial impulse. The expectation below reflects a sophisticated agent's beliefs about temptation-phase impulses.

$$V^s(m, I) = \max_{c^1, c^2 \in m} \mathbb{E}^s_{(m)}[U(c^1, c^2; I, I')|I] \tag{4.2}$$

where $\mathbb{E}_{(m)}$ is the expectation with respect to the distribution $\Phi'(m)$ of I'.

The consumer chooses a menu m from menu of menus $\mathcal{M}$ and consumption from that menu, given an initial impulse I, expected temptation-phase impulses, and prices. Gul and Pesendorfer (2001) provide a utility function representation that allows for ranking of menus, based on how consumers evaluate choices before and after temptation is manifest. Since the impulses in my model are stochastic, the consumer can only anticipate temptations and decide based on expected utility. The consumer's problem in a one-period model is to choose a menu that attains the maximum expected surplus, which is given by

$$\max_{m \in \mathcal{M}} V^s(m, I) - p_m \tag{4.3}$$

4.3.2 Behavioral Types

Consumers differ in their knowledge of own self-control and susceptibility to addiction. *Sophisticated* consumers are aware of the difference in strength between temptation-phase impulses with and without a menu that includes more than one unit of the good. Consumers who are unaware of this difference are *naive* and behave as if according to a different value function, V^n. They choose a menu m to solve:

$$\max_{m \in \mathcal{M}} V^n(m, I) - p_m \tag{4.4}$$

The value function $V^n(m, I)$ obtains from maximization of expected utility, as V^s does, but with respect to a mistaken distribution that underestimates the probability of non-negligible impulses. An extreme case of naivete would be the belief that temptation-phase impulses are drawn from $\Phi'(\emptyset)$ when they are in fact drawn from $\Phi'(L)$.

4.3.3 Consumers in the Multi-Period Model

In each time period, consumers make decisions about which menu to purchase and how much to consume. In addition to pecuniary and health costs and the discomfort of unfulfilled impulses, consumers must also weigh the risks of addiction. Addiction is represented by an increased likelihood of non-negligible impulses in future periods following consumption. Each period's initial-phase impulses are drawn from a distribution Φ_t, which evolves from one period to the next, as a function of the level of consumption in the previous period, $c_{t-1} = c^1_{t-1} + c^2_{t-1}$.

Assumption 12. *The evolution of initial-phase impulses is such that Φ_{t+1} first-order stochastically dominates Φ_t if $c_t \geq 1$.*

The timeline below summarizes impulses and decisions within each time period and adds the evolution of initial-phase impulses described in Assumption 12.

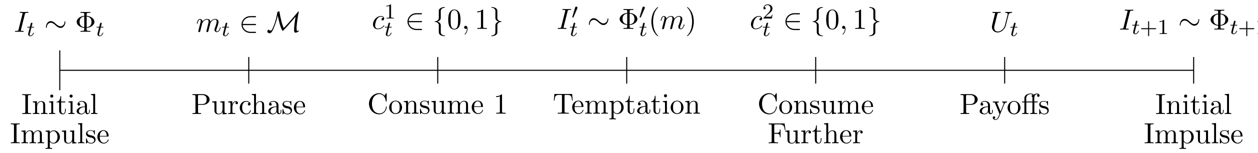

Figure 4.3: Timing of impulses and decisions in the multi-period model

In each period, the utility realized depends on consumption and impulses.

$$U_t(c_t^1, c_t^2; I_t, I_t') = (1 - c_t^1)(-I_t) + (1 - c_t^2)(I_t') - H(c_t^1 + c_t^2) \tag{4.5}$$

In the one-period model, the value function specifies the maximum utility that could be obtained, given a menu and initial-phase impulse; the consumer's problem is to choose the menu that maximizes surplus: utility minus cost. In the multi-period setting, there is an interdependence between one-period value functions and spending on menus. The menu chosen in the present period affects consumption, which in turn affects the next period's impulses and menu choices. Hence, the multi-period value function will describe the maximum utility obtainable, given a series of menus and an initial impulse. The consumer's problem is then to choose the series of menus that maximizes time-discounted surplus. For simplicity, prices of menus are assumed to be time-independent. If $m_t = m_{t+1}$ then $p_{m_t} = p_{m_{t+1}}$.

The formulation below is recursive, in that the present period's value depends on the next period's value. Sophisticated consumers choose c_t^1 and c_t^2 in period t knowing their effect on the distribution of impulses in period $t+1$. The multi-period value function is

$$V_t^s(\{m_k\}_{k=t}^T, I_t) = \max_{(c_t^1, c_t^2) \in m_t} \mathbb{E}^s_{(m_t)}[U(c_t^1, c_t^2; I_t, I_t') + \beta V_{t+1}^s(\{m_k\}_{k=t+1}^T, I_{t+1}) | I_t] \tag{4.6}$$

where I_{t+1} is distributed according to Φ_{t+1} which depends on Φ_t and c_t.

The consumer's problem is to choose the sequence of menus that solves

$$\max_{\{m_k\}_{k=t}^{T}} V_t^s(\{m_k\}_{k=t}^{T}, I_t) - \sum_{k=t}^{T} \beta^{k-t} p_{m_t} \tag{4.7}$$

where $V_{T+1}^s = 0$.

4.3.4 Behavioral Types in the Multi-Period Model

Consumers may differ in knowledge of their own vulnerabilities to addiction. A *farsighted* consumer predicts, with complete accuracy, the effect that present consumption will have on the future impulse distribution; a *myopic* consumer underestimates this addictive effect of consumption. Myopia and farsightedness can interact with awareness of the effect of temptation in the short-term. The *sophisticated, farsighted* consumer is aware of both the temptation that will occur and the effect of consumption on future urges, and plans for those. The *sophisticated, myopic* consumer knows the temptation that will occur, and plans for it, but is unaware of the effect of consumption on future impulses. We can assign some mistaken impulse evolution process $\widetilde{\Phi}_t$ which depends on Φ_{t-1} and c_{t-1}.

The *naive, farsighted* consumer understands the effect of consumption on future urges but is unaware of his self-control problem in the face of temptation. Assign a mistaken value of $\widetilde{\Phi}'_t(L)$ to represent the naive consumer's forecasted temptation-phase beliefs following purchase of a large menu. In the extreme, $\widetilde{\Phi}'_t(L) = \widetilde{\Phi}'_t(\emptyset)$. The *naive, myopic* consumer is aware of neither his self-control problem nor the effect of consumption on future urges. This is often a first-time consumer with no education on these post-consumption effects.

4.3.5 Policy

Welfare maximization may be brought about by reducing consumption of certain harmful goods that are both tempting and addictive. If the government has evidence that consumption of a harmful good is often just a response to addiction, the policy may include provisions to prevent inexperienced consumers from trying the good, such as prohibition of small menus. On the other hand, if the temptation following purchase of large menus is overwhelming, the government may wish to implement a policy that makes small menus available. The analysis that follows assumes that consumers may always choose either an empty menu or a large menu, but compares the choices with and without the availability of a small menu.

4.4 Model with Discrete Impulse Distribution

To simplify the analysis and develop intuition, I consider a model that discretizes the impulse distribution to three states. Government policy or firms' choices may remove the availability of S. The main question is to identify conditions under which removal of S may improve welfare.

At the beginning of each time period, an impulse, I, is realized, which can be either intense (I_T), weak (I_W), or negligible (I_N), with probabilities $\mathbb{P}(T)$, $\mathbb{P}(W)$, and $\mathbb{P}(N)$ that sum to 1. A negligible impulse is one that the consumer can coldly reject without exercising self-control, while weak and intense impulses cause discomfort if left unfulfilled. The decision-maker can choose to purchase either $\emptyset$, L, or, if it is available, S, with prices $p_L > p_S$.[9] Additionally, consumption of each unit of the good carries a health cost of H.

The consumer is willing to purchase the large or small menu and suffer a health cost to relieve the intense impulse, but when faced with a weak impulse, is only willing to purchase the small menu. Expected temptation-phase impulses may also enter the decision-making process. If the consumer has purchased either a small or large menu, he immediately consumes one unit of the good to fulfill a weak or intense impulse.

Assumption 13. *Willingness to remove the intense impulse is greater than the cost of the large menu, and willingness to remove the weak impulse is strictly between the costs of the*

[9] We may assume that $2p_S > p_L$. Otherwise, there is no point in offering L.

large menu and the small menu. That is:

$$I_T - H > p_L > I_W - H > p_S$$

This relationship serves to discretize the impulse distribution described in section 4.3: If $I > p_L + H$, it is classified as an intense impulse, and the probability of this is the probability that a draw from Φ is greater than $p_L + H$. Likewise, if $I \in (p_S + H, p_L + H]$, it is classified as a weak impulse. The prices of the menus are assumed to be exogenous but if one wishes to model the decision of the firm this assumption can be modified. I assume that p_L and p_S account for transaction costs.

After the initial purchase and consumption decisions, a consumer is faced with the temptation to consume further, in the form of a second impulse, I'. If either the small or empty menus were purchased, this temptation-phase impulse will go unfulfilled. The large menu allows for impulse fulfillment in the temptation phase, but *creates* additional temptation.[10]

The probability of non-negligible impulses in the temptation phase depends on the menu the consumer purchased. Let $\mathbb{P}_R$ give the probability of each impulse in the temptation phase if there are unconsumed items remaining, and $\mathbb{P}_\emptyset$ the probability of each type of impulse if there is no possibility of consumption.[11] The difference $[\mathbb{P}_R(T) + \mathbb{P}_R(W)] - [\mathbb{P}_\emptyset(T) + \mathbb{P}_\emptyset(W)]$ is one measure of the consumer's self-control. A higher value indicates lower self-control.

Assumption 14 adapts Assumption 11 to the discrete model. As impulses with higher values become more likely, they become more likely to be classified as intense, and more likely to be classified as non-negligible.

Assumption 14. *(Temptation)* $\mathbb{P}_R(T) \geq \mathbb{P}_\emptyset(T)$ *and* $\mathbb{P}_R(T) + \mathbb{P}_R(W) \geq \mathbb{P}_\emptyset(T) + \mathbb{P}_\emptyset(W)$.

[10] Note also the marginal pecuniary cost of additional consumption is zero.

[11] If S is purchased and one unit consumed, temptation-phase impulses will be unfulfilled. The same goes for the case where the consumer purchased $\emptyset$.

After purchase, consumption, and savings decisions are made, payoffs are tallied. If the consumer has chosen menu m from $\{\emptyset, S, L\}$, and has chosen consumption in initial and temptation phases, c^1 and c^2, each from $\{0, 1\}$ then utility is

$$U(c^1, c^2; I, I') = -(1 - c^1)(I) - (1 - c^2)(I') - (c^1 + c^2)H \tag{4.8}$$

and surplus is $U(c^1, c^2; I, I') - p_m$. The consumer chooses the menu m and, given the temptation-phase impulses that follow, the consumption c^1 and c^2 that maximizes expected surplus, $\mathbb{E}^s_{(m)}[U(c^1, c^2; I, I')|I] - p_m$.

If the consumer purchased the large menu, it may have been with a plan to coldly satisfy the expected impulses according to $\mathbb{P}_\emptyset$. However, with the actual temptation of the large menu, impulses occur more frequently according to $\mathbb{P}_R$, and when fulfilled do not increase the payoff to the long-run self. $\mathbb{P}_R$ also implies a higher expected health cost relative to $\mathbb{P}_\emptyset$. The expected payoff in the temptation phase following the purchase of a large menu, initial consumption of one unit, and additional consumption conditional upon non-negligible impulse is $-H[\mathbb{P}_R(T) + \mathbb{P}_R(W)]$ and the total expected utility is

$$\mathbb{E}^s_{(L)}[U(1, c^2, I, I')] = -H - H[\mathbb{P}_R(T) + \mathbb{P}_R(W)] \tag{4.9}$$

I assume that if the consumer is able to resist an initial impulse, temptation-phase impulses alone are insufficient to motivate purchase of even the small menu. Note that the consumer that plans for temptation-phase impulses by purchasing the good and waiting to consume it would face impulses from $\mathbb{P}_R$.

Assumption 15. *Expected impulses in the temptation phase alone are not sufficient motivation for the purchase of a small menu. That is,* $p_S > \mathbb{P}_R(T)(I_T - H) + \mathbb{P}_R(W)(I_W - H)$.

Due to a commitment to the small menu there may be unfulfilled impulses in the temptation phase.[12] These impulses will not occur as frequently as when the consumer has purchased

[12] We could allow the consumer to make an additional purchase at this point but I do not study this situation here.

the large menu. The expected payoff in the temptation phase following the purchase of a small menu is $-\mathbb{P}_\emptyset(T)I_T - \mathbb{P}_\emptyset(W)I_W$ and the total expected utility, including the health cost of initial consumption is

$$\mathbb{E}^s_{(S)}[U(1,0,I,I')] = -H - \mathbb{P}_\emptyset(T)I_T - \mathbb{P}_\emptyset(W)I_W \tag{4.10}$$

The empty menu commits the consumer to abstention and the unfulfilled impulses that accompany it. The total expected utility is

$$\mathbb{E}^s_{(\emptyset)}[U(0,0,I,I')] = -I - \mathbb{P}_\emptyset(T)I_T - \mathbb{P}_\emptyset(W)I_W \tag{4.11}$$

4.4.1 Decisions in the One-Period Model

In the one-period model addiction is not a factor; the consumer simply chooses the menu m which gives the greatest expected surplus, $V^s(m,I) - p_m$. Consumers are risk-neutral with respect to temptation-phase impulses and must consider how purchase of the large menu changes the impulse distribution in the temptation phase. The optimal decision rule for the consumer depends on the prices of the small and large menus, p_S and p_L, as well as consumer-specific parameters: health costs H, the strength of the initial impulse (I_T or I_W), and the distribution of temptation-phase impulses ($\mathbb{P}_\emptyset$ and $\mathbb{P}_R$).

Lemma 15. *A sophisticated consumer's optimal decision rule when faced with a menu of menus* $\{\emptyset, S, L\}$ *is:*

1. *In the case of intense impulse* I_T*: Choose* L *if* $p_L + H[\mathbb{P}_R(T) + \mathbb{P}_R(W)] \leq p_S + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W$ *and choose* S *otherwise.*

2. *In the case of weak impulse* I_W*: Same decision rule.*

3. *In the case of negligible impulse* I_N*: Choose* $\emptyset$*.*

Proof. In cases 1 and 2, the empty menu is dominated by the small menu. It would be better to choose $\emptyset$ only if $I_T \leq p_S + H$ but this violates Assumption 13. The payoffs in

(4.9) and (4.10), net of the prices of the menus, form the basis for comparison of the large and small menus. Case 3 is an immediate consequence of Assumption 15. □

Lemma 16. *A sophisticated consumer's optimal decision rule, when faced with a menu of menus* $\{\emptyset, L\}$ *is:*

1. *In the case of intense impulse* I_T*:*
   Choose L if $p_L + H[1 + \mathbb{P}_R(T) + \mathbb{P}_R(W)] \leq I_T + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W$ *and choose* $\emptyset$ *otherwise.*

2. *In the case of weak impulse* I_W*:*
   Choose L if $p_L + H[1 + \mathbb{P}_R(T) + \mathbb{P}_R(W)] \leq I_W + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W$ *and choose* $\emptyset$ *otherwise.*

3. *In the case of negligible impulse* I_N*: Choose* $\emptyset$*.*

Proof. The payoffs in (4.9) and (4.11), net of the prices of the respective menus, form the basis for comparison of the large and empty menus. □

Proposition 17. *The relationship between policy on available menus and consumer choice is as follows:*

1. *Any sophisticated consumer that chooses L from* $\{\emptyset, L\}$ *would choose L or S from* $\{\emptyset, S, L\}$*.*

2. *Any sophisticated consumer that chooses L from menu of menus* $\{\emptyset, S, L\}$ *would also choose L from* $\{\emptyset, L\}$*.*

Proof. (1) When only menu L is available and an intense impulse is realized, the consumer chooses menu L if (4.13) holds. Following a weak impulse, the consumer chooses L if (4.14) holds. The condition for choosing a non-empty menu, S or L, is simply Assumption 13, $p_S < I_T - H$. Either of (4.13) or (4.14) represents an additional necesary condition for purchase of a non-empty menu when choices are restricted to $\{\emptyset, L\}$.

(2) When menu S is available and either an intense or weak impulse is realized, the consumer chooses menu L if

$$p_L + H[\mathbb{P}_R(T) + \mathbb{P}_R(W)] \leq p_S + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W \tag{4.12}$$

and chooses S otherwise.

When menu S is not available, the consumer chooses menu L following an intense impulse if

$$p_L + H[\mathbb{P}_R(T) + \mathbb{P}_R(W)] \leq I_T - H + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W \tag{4.13}$$

and chooses menu L following a weak impulse if

$$p_L + H[\mathbb{P}_R(T) + \mathbb{P}_R(W)] \leq I_W - H + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W \tag{4.14}$$

According to Assumption 13, the right-hand side of both (4.13) and (4.14) is larger than the right-hand side of (4.12). Hence, any consumer that chose L when S was available will also choose L when S is not available, but the converse is not true. Some consumers will choose S instead of L when both menus are available. □

This result tells us that when choices of menus are restricted to exclude the small menu, consumption of compulsive goods decreases at the extensive margin – more people choose the empty menu – and increases in the intensive margin – of those who give in to the impulse, more of them purchase the large menu. Because of the nature of temptation, we would expect this to bring about greater consumption among those who purchase the good.

Effect of Policy

Absent any restrictions on the menus that are available, consumers rank the three menus $\emptyset$, S, and L according to the expected payoffs in (4.9), (4.10), and (4.11). Preference rankings arise from the strength of the initial impulse, the relative prices of the small and large menus, health costs, and the probabilities of temptation-phase impulses.

Some consumers will have strong and frequent cravings, corresponding to preference ranking $L > S > \emptyset$.[13] The policy of removing S from the choice set will not induce abstention. If a consumer's ranking is $\emptyset > S > L$,[14] the consumer faces negligible impulses, and is not made worse off by the presence of the small menu.

It is the consumer for whom $S > \emptyset > L$ [15] that "large menu or none" policies target. Removing the small menu option induces these consumers to abstain from purchase of the unhealthy product. Since our analysis of the one-period model ignores addiction, the implication of a policy that induces the choice of a second-best option is that the government knows the consumer is better off net of factors such as addiction and mistaken hedonic forecasting which are not in the one-period sophisticated agent model.

In the case where $S > L > \emptyset$, the presence of the small menu improves the welfare of a consumer who would otherwise choose the large menu to fulfill an intense impulse; the small menu option also allows for lower consumption of the unhealthy good. These consumers would be hurt by a policy restricting the sale of tempting goods to only large menus.

4.4.2 Consumer Naivete

Some consumers are sophisticated and know the extent of their self-control problems, while others are naive and incorrectly estimate the effect of temptation. A naive consumer chooses a menu anticipating impulses from $\mathbb{P}_\emptyset$ but a sophisticated consumer knows better and anticipates the increased temptation of $\mathbb{P}_R$. If a naive consumer perceives that temptation-phase impulses are distributed according to $\mathbb{P}_\emptyset$ when in reality they are distributed according to $\mathbb{P}_R$, the evaluation of (4.9) is adjusted accordingly and a large menu may be chosen. Ex-post, when faced with temptation, the consumer may regret her choice of menu.

[13] $L > \emptyset > S$ could also be included in this category but perhaps it is pathological.

[14] Same for $\emptyset > L > S$. The small menu option is inconsequential, and these preferences seem pathological.

[15] See lemmas 25 and 26 in the appendix for conditions under which these preferences arise.

Lemma 18. *The naive consumer's optimal decision rule when faced with menu of menus* $\{\emptyset, S, L\}$ *and either* I_T *or* I_W *is*

Choose L if $p_L + H[\mathbb{P}_\emptyset(T) + \mathbb{P}_\emptyset(W)] \leq p_S + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W$ *and choose S otherwise.*

Lemma 19. *The naive consumer's optimal decision rules when faced with menu of menus* $\{\emptyset, L\}$ *are:*

1. *In the case of intense impulse* I_T*:*
   Choose L if $p_L + H[1 + \mathbb{P}_\emptyset(T) + \mathbb{P}_\emptyset(W)] \leq I_T + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W$ *and choose* $\emptyset$ *otherwise.*

2. *In the case of weak impulse* I_W*:*
   Choose L if $p_L + H[1 + \mathbb{P}_\emptyset(T) + \mathbb{P}_\emptyset(W)] \leq I_W + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W$ *and choose* $\emptyset$ *otherwise.*

The analysis in Section 4.4.1 implies that the policy of removing the small menu from the choice set cannot improve the welfare of a sophisticated consumer. It is possible that this policy can protect a naive consumer but the following result combines Lemmas 18 and 19 to show that limiting menus to $\{\emptyset, L\}$ does not change the effect of naivete in the one period model. We can re-write the decision rules for sophisticated and naive consumers to find the maximum value of p_L that would lead to purchase of the large menu, holding p_S and the consumer parameters fixed.

Proposition 20.

1. *If* $\emptyset$*, S, and L are all available, a naive consumer with* $\mathbb{P}_R > \mathbb{P}_\emptyset$ *has a higher willingness to pay for L than a sophisticated consumer.*

2. *If only* $\emptyset$ *and L are available, a naive consumer has a higher willingness to pay for L than a sophisticated consumer.*

Proof. (1) Rewriting the decision rules in lemmas 15 and 18, and holding all other parameters constant, we can find upper bounds for p_L that would induce naive and sophisticated consumers to choose L over S when faced with either a weak or intense impulse.

The upper bound on p_L for a naive consumer exceeds the upper bound for a sophisticated consumer by $H[\mathbb{P}_R(T) + \mathbb{P}_R(W) - (\mathbb{P}_\emptyset(T) + \mathbb{P}_\emptyset(W))]$. This wedge is positive since $\mathbb{P}_R(T) + \mathbb{P}_R(W) > \mathbb{P}_\emptyset(T) + \mathbb{P}_\emptyset(W)$.

(2) Rewriting the decision rules in lemmas 16 and 19, we can find upper bounds for p_L that would induce naive and sophisticated consumers to choose L over $\emptyset$. This generates the same result that the upper bound for a naive consumer exceeds the upper bound for a sophisticated consumer by $H[\mathbb{P}_R(T) + \mathbb{P}_R(W) - (\mathbb{P}_\emptyset(T) + \mathbb{P}_\emptyset(W))]$. □

Proposition 20 shows that the distortionary effect of naivete on the perceived value of the large menu is not mitigated by removal of the small menu. It is also useful to notice that naivete about temptation-phase impulses does not affect the payoffs to consumers who choose the small menu; expectations of $\mathbb{P}_\emptyset$ following purchase of S would be correct.

4.4.3 Two-Period Model

After all decisions and accounting for the present period are finished, the consumer's impulse distribution evolves for the next period. To bring addiction into the model, allow the probabilities of intense and weak impulses to increase when at least one unit of the good is consumed in period 1, consistent with Assumption 12. I use $\mathbb{P}(T_1)$ and $\mathbb{P}(W_1)$ for the probabilities of I_T and I_W in period 1, and $\mathbb{P}(T_2|c_1)$ and $\mathbb{P}(W_2|c_1)$ for the probabilities of intense and weak impulses in period 2, conditional upon total period-1 consumption, $c_1 = c_1^1 + c_1^2$.

Assumption 16. *(Addiction) For* $c_1 \geq 1$, $\mathbb{P}(T_2|c_1) \geq \mathbb{P}(T_1)$ *and* $\mathbb{P}(T_2|c_1) + \mathbb{P}(W_2|c_1) \geq \mathbb{P}(T_1) + \mathbb{P}(W_1)$.

Additionally, it may be assumed that addiction is increasing in consumption: $\mathbb{P}(T_2|c_1 = 2) \geq \mathbb{P}(T_2|c_1 = 1)$ and $\mathbb{P}(T_2|c_1 = 2) + \mathbb{P}(W_2|c_1 = 2) \geq \mathbb{P}(T_2|c_1 = 1) + \mathbb{P}(W_2|c_1 = 1)$.

In this paper, I do not allow the consumer to save storable tempting goods, but we can reflect on the consequences of adding this feature to the model. If there are saved goods on hand and the replacement cost is stationary, the consumer facing a weak impulse will fulfill it by consumption of the saved goods. It is even possible that a negligible impulse may be fulfilled, but according to the model specified above, if the consumer rejected a temptation-phase impulse to save the remainder of the menu, then he can reject the negligible impulse.

I examine the benchmark case of the sophisticated, farsighted consumer who knows the distribution of temptation-phase impulses and the evolution of the impulse distribution, which represents the addiction process. Optimal actions can be determined by backward induction. In the second period, future impulses due to addiction are not a factor, but we can still allow for costs to be as before, where the consumer coldly rejects a neglible impulse. Lemmas 15 and 16 provide the decision rules for period 2.

Period 2 payoffs are determined by prices p_S and p_L and consumer parameters H, I_T, I_W which are stationary across time, a discount factor $\beta \in (0,1]$, the probabilities of intense and weak impulses in the initial phase of each period, $\mathbb{P}(T_1)$, $\mathbb{P}(W_1)$, $\mathbb{P}(T_2|c_1)$, $\mathbb{P}(W_2|c_1)$, and the probabilities of intense and weak impulses in the temptation phase of each period, $\mathbb{P}_\emptyset(T_1)$, $\mathbb{P}_\emptyset(W_1)$, $\mathbb{P}_R(T_1)$, $\mathbb{P}_R(W_1)$, $\mathbb{P}_\emptyset(T_2)$, $\mathbb{P}_\emptyset(W_2)$, $\mathbb{P}_R(T_2)$, $\mathbb{P}_R(W_2)$. The set of period-2 temptation-phase probabilities are dependent on both period-2 menu choice and period-1 consumption.

Some reasonable simplifying assumptions can be made to reduce the dimensionality of the problem. In keeping with Assumption 16, consumption in the present period makes the intense impulse more likely in the next period. With $\hat{K} \geq K \geq 1 \geq k$,

$$\mathbb{P}(T_2|c_1) = \begin{cases} \hat{K}\mathbb{P}(T_1), & \text{if } c_1 > 1 \\ K\mathbb{P}(T_1), & \text{if } c_1 = 1 \\ k\mathbb{P}(T_1), & \text{if } c_1 = 0 \end{cases} \tag{4.15}$$

and the same relationship holds for $\mathbb{P}(W_1)$ and $\mathbb{P}(W_2)$.[16]

[16] This assumes that $\hat{K}(\mathbb{P}(T_1)+\mathbb{P}(W_1)) \leq 1$. If not, then increases in $\mathbb{P}(T)$ would have to be accompanied

Let the probabilities of temptation-phase impulses be a constant multiple of the initial impulses in their respective periods.

Assumption 17. *(1)* $\mathbb{P}_\emptyset(T_t) = a_\emptyset \mathbb{P}(T_t)$ *and* $\mathbb{P}_\emptyset(W_t) = a_\emptyset \mathbb{P}(W_t)$ *for* $t = 1, 2$. *(2)* $\mathbb{P}_R(T_t) = a_R \mathbb{P}(T_t)$ *and* $\mathbb{P}_R(W_t) = a_R \mathbb{P}(W_t)$ *for* $t = 1, 2$. *(3)* $a_\emptyset \in (0, 1)$ *and* $a_R \in (a_\emptyset, 1)$.

Combining this assumption with the relationships described in (4.15), all impulse probabilities can be related by constant multipliers to the period 1 initial-phase impulses. In the first period, the sophisticated consumer knows the decision rules for period 2, and uses this knowledge to weigh the period 1 impulses against health costs and effects of addiction. Due to the many parameters, the problem is not highly tractible, but I will walk through the steps of analysis and provide some general results. The following two lemmas describe the decision rules for the consumer in period 2.

Lemma 21. *Under Assumption 17, the optimal decision rules for period 2 when faced with a menu of menus* $\{\emptyset, S, L\}$ *are:*

1. *In the case of intense impulse* I_T*:*
   Choose L *if* $p_L + Ha_R[\mathbb{P}(T_2|c_1) + \mathbb{P}(W_2|c_1)] \leq p_S + a_\emptyset[\mathbb{P}(T_2|c_1)I_T + \mathbb{P}(W_2|c_1)I_W]$ *and choose* S *otherwise.*

2. *In the case of weak impulse* I_W*: Same decision rule.*

3. *In the case of negligible impulse* I_N*: Choose* $\emptyset$.

Proof. These are the conditions from Lemma 15 rewritten using (4.15) and Assumption 17. □

Lemma 22. *Under Assumption 17, the optimal decision rules for period 2 when faced with a menu of menus* $\{\emptyset, L\}$ *are:*

by decreases in $\mathbb{P}(W)$ or restricted by some upper bound.

1. *In the case of intense impulse* I_T*:*
   Choose L *if* $p_L + H + Ha_R[\mathbb{P}(T_2|c_1) + \mathbb{P}(W_2|c_1)] \leq I_T + \emptyset[\mathbb{P}(T_2|c_1)I_T + \mathbb{P}(W_2|c_1)I_W]$ *and choose* $\emptyset$ *otherwise.*

2. *In the case of weak impulse* I_W*:*
   Choose L *if* $p_L + H + Ha_R[\mathbb{P}(T_2|c_1) + \mathbb{P}(W_2|c_1)] \leq I_W + \emptyset[\mathbb{P}(T_2|c_1)I_T + \mathbb{P}(W_2|c_1)I_W]$ *and choose* $\emptyset$ *otherwise.*

3. *In the case of negligible impulse* I_N*: Choose* $\emptyset$*.*

Proof. These are the conditions from Lemma 16 rewritten using (4.15) and Assumption 17. □

Period 1 begins with either I_T, I_W, or I_N. Though Assumption 15 precludes the possibility of doing so in period 2, it is possible for the consumer to purchase the large menu in period 1, then resist the subsequent temptation. Clearly, S is strictly preferred to buying L then resisting temptation, though if S is unavailable, buying L then resisting either the weak impulse or both weak and intense impulses may be preferred to the empty menu.[17] The expected payoffs for each period-1 menu and consumption plan are provided in Section C.2 of the appendix.

$V^2(\cdot)$ is the period-2 "value" of following the optimal decision rule in period 2, with the addiction update factor, $\hat{K}$, K, or k, as its argument.[18] Conveniently, because of the relationship described in (4.15), the values of $V^2(k)$, $V^2(K)$, and $V^2(\hat{K})$ differ only by the constant factors relating the period 1 and period 2 initial-phase impulse distributions.

$$\begin{aligned} V^2(K) &= \frac{K}{k}V^2(k) \\ V^2(\hat{K}) &= \frac{\hat{K}}{k}V^2(k) \end{aligned} \tag{4.16}$$

[17] Purchasing S dominates purchasing L then not consuming in the temptation phase. However, the latter may be the best option if consumers can save for period 2.

[18] See appendix B for full expansion of the period-2 value.

When facing a negligible impulse, the consumer will choose $\emptyset$. This is because period-1 impulses are not strong enough to justify purchase, and choice of the empty menu will minimize the probability of impulses in period 2. Other definitive rules are difficult to pin down, but Lemma 23 summarizes some facts about how the parameters of the problem affect the optimal period-1 decision when a consumer faces a non-negligible impulse.

Lemma 23. *For a given consumer, the parameters of the two-period model affect the optimal period-1 decision in the following ways:*

(1) An increase in H lowers the utility of S and L and makes $\emptyset$ relatively more attractive.

(2) An increase in prices makes $\emptyset$ more attractive and an increase in p_L/p_S increases the utility of S relative to L.

(3) An increase in $\frac{a_R}{a_\emptyset}$ increases the utility of choices that reduce the effect of temptation, S and $\emptyset$, relative to L.

(4) Decreases in $\hat{K}$ and K, the effects of addiction, increase the utility of L and S relative to $\emptyset$.

(5) The effects of I_T and I_W are ambiguous and depend on the tradeoffs between period-1 and period-2 disutility, which is in turn dependent on β and the effects of addiction.

4.5 Future Extensions

The model described in this paper assumes that menu prices are exogenous, and policy tools are limited to restriction of the small menu. In reality, firms choose menu sizes, prices, and positioning of products among other things, and policy may provide guidance on these and other decisions. Firms that have done market research know approximately how many consumers are in each of the categories of behavioral types mentioned in Section 4.3.4. In order to maximize their profits, firms choose a menu of menus in the first period, with commitment, knowing the characteristics of consumers and their likely decisions. Government policy may restrict the firms' choices.

Cigarette manufacturers in the U.S. currently sell cigarettes in packs of twenty and cartons of ten packs. Convenience stores sell small packages of candy and cookies to fulfill impulses. Supermarkets sell both small and large packages, though the small packages are strategically positioned where consumers linger, in order to *induce* impulses rather than fulfill existing impulses.

It is not clear what the overall effect of small menus is on total profit. Small menus can help consumers avoid excess temptation that leads to addiction and help casual consumers with their self-control problems. As discussed in section 4.4.1, some consumers who would otherwise select a small menu may rather purchase a large menu than nothing and if follow-on temptation is consistent with Assumption 14, this can lead to higher-than-intended consumption and addiction. Creating temptation and addiction may be bad for public relations, but the firm can counter by shifting responsibility to the consumer. Some food companies depict a culture of sports, exercise, and movement, sending the message that weight problems are a result of lagging personal responsibility and insufficient exercise, rather than overconsumption triggered by advertising and the changing composition of foods (Brownell and Warner, 2009).

Retailers may wish to prioritize sales of packs to generate higher volume of sales, and if offered, to price single cigarettes at a steep markup as an altenative for consumers with occasional impulses. Secondhand vendors have emerged in many market to divide otherwise indivisible goods. Businesses exist to help people share cars, apartments, and many other durable goods. Yet among consumer-packaged goods, analogous "sharing" or division of a large package into single units is often prohibited by restrictions put in place by manufacturers and governments. Removing the restriction on the sale of single cigarettes would allow retail outlets to capture an additional segment of the market, but this may be a nuisance for them. Still, decrimininalizing distribution by ordinary citizens or allowing a limited number of licenses makes the market contestable and incentivizes the retailer to offer single cigarettes for a competitive price.

4.6 Conclusion

Depending on the distribution of consumer types in a population, it may be welfare-improving to offer goods in a variety of package sizes to consumers, so impulses may be fulfilled while minimizing temptation and addiction. Even if the price of a single unit of a compulsive good is close to the price of a multi-unit pack, the availability of the option can reduce harmful excess consumption. This may be relevant to policy on minimum cigarette package size.

Another application is in the arena of foods (Brownell and Warner, 2009), specifically those foods for which consumers have impulses. A small serving can fulfill an impulse, but a package containing multiple servings can lead to excess temptation and consumption, which in turn perpetuates future impulses. With foods which consumers are unlikely to view as harmful or potentially addictive, there is a strong heuristically-driven motive to seek the best bargain. Consumers may see the high per-unit price of addictive foods in small packages as “unfair” or “uneconomical” and opt for a package that is larger than the impulse demanded. Once in possession of the larger package, it is easy for the consumer, naive about the health and addiction costs, to consume more than planned. This suggests that pricing guidance may be a welfare-improving policy.

There is an important caveat to promoting a wider variety of package sizes as a harm reduction strategy. Smaller portion packages can lead to greater future consumption among those who are not already addicted or experienced in use of the product. For the type of consumer who has weak impulses, a small package provides cost-effective gratification, and no consumption is a second-best alternative. If this type of consumer is also naive about their own addiction, then it may be best not to allow access to the smaller package and disincentivize purchase by consumers who are merely curious or have occasional cravings. Empirical studies on the strength of impulses and likelihood of addiction following infequent consumption would be necessary to address this question.

Chapter 5

Conclusion and Discussion

Social dissonance games can be used to model situations where the conformity motive plays a role in individual decision-making. My analysis focuses on games with two actions. In the finite-agent games discussed in the first chapter, equilibria have a monotonicity property: those players who intrinsically like an action the most will be the ones to choose it. These games are supermodular, which is to say there are complementarities among the players' strategies. By applying the results of Milgrom and Roberts (1990) we can guarantee existence of a pure strategy equilibrium and monotone comparative statics with respect to intrinsic utilities and social dissonance. If the intrinsic utility of one action relative to the other increases for all players, the equilibrium set shifts toward action profiles where more players choose that action. Likewise, if the social dissonance for one action decreases relative to the other action for all possible profiles, the equilibrium set again shifts toward that action.

Analogous results hold in social dissonance games with a continuum of agents, as discussed in the second chapter. Equilibria are monotone and the games are supermodular, hence the results of Yang and Qi (2013) can be used to guarantee monotone comparative statics with respect to intrinsic utilities and social dissonance. However, the finite-agent and continuum-of-agents models diverge with respect to the asymmetric action profiles that may be equilibria. In finite-agent games with an assumption that dissonance is strictly decreasing, it is possible to have asymmetric behavior among a group of agents who are indifferent between two alternatives, but any two identical agents that strictly prefer one of the alternatives must choose alike. In games with a continuum of agents, asymmetric behavior may be observed in equilibrium among a positive-measure group of agents with the same intrinsic preferences.

While asymmetric equilibria may seem like razor-edge scenarios, they can be important in understanding the preservation of inefficient norms. In the finite agent games, coordination of indifferent agents on a single action can cause unnecessary discomfort for those who prefer the other action, relative to an equilibrium where these indifferent agents are more evenly split between the two actions. In the continuum of agents game with players divided into two camps each with identical intrinsic preferences, one camp may be split into those who are true and those who sell out.

Just as the actions and beliefs of people around us, actual or imagined, influence our own behavior, the objects we purchase can also influence our behavior in ways that we do not anticipate. In "Impulse and Temptation," I set up a framework for answering the question of whether harmful consumer goods should be available in small packages to satisfy cravings. Policy must balance the lower cost of the small package, which lowers the threshold of a craving that induces purchase, against the increased temptation and subsequent habit formation associated with the larger package. In a version of the model with a discretized distribution of impulses to consume the harmful good, the effect of making the small package unavailable depends on the second-best alternative for consumers who prefer the small package. In general, fewer consumers will purchase the product, but of those who do, more will purchase the large package and be tempted toward greater consumption.

Since a policy banning the sale of small packages of harmful goods seems welfare-reducing in a one-period model, I explore alternative justifications for the policy. Consumer naivete about temptations is one possible justification. However, removing the small menu from consideration does nothing to alleviate mistakes due to incorrectly forecast temptations. In a two-period model, it is unclear whether prohibiting the sale of small menus helps consumers to avoid addiction or improve welfare. This question would require data on the perceived and actual likelihood of addiction following the consumption of small quantities of impulse goods.

References

[1] Acemoglu, Daron and Martin Kaae Jensen (2010). “Robust Comparative Statics in Large Static Games.” 49th IEEE Conference on Decision and Control (CDC), 3133-39

[2] Akerlof, George A. (1980): “A Theory of Social Custom, of Which Unemployment May be One Consequence.” Quarterly Journal of Economics, 94(4).

[3] Asch, Solomon E. (1951): “Effects of Group Pressure Upon the Modification and Distortion of Judgments.” *Groups, leadership and men; research in human relations.* Ed. Guetzkow, Harold. Oxford, England: Carnegie Press, ix, pp. 177-190.

[4] Athreya, Kartik (2004): “Shame As It Ever Was: Stigma and Personal Bankruptcy.” Federal Reserve Bank of Richmond Economic Quarterly 90/2.

[5] Aumann, Robert J. (1964). “Markets with a Continuum of Traders.” Econometrica, vol. 32, 39-50.

[6] Banerjee, Abhijit (1992): “A Simple Model of Herd Behavior.” Quarterly Journal of Economics, 107(3).

[7] Becker, Gary and Kevin Murphy (1988). “A Theory of Rational Addiction.” Journal of Political Economy, 96:4, p. 675-700.

[8] Becker, Gary, Michael Grossman and Kevin Murphy (1994). “An Empirical Analysis of Cigarette Addiction.” American Economic Review, 84:3, p. 396-418.

[9] Bernheim, B. Douglas (1994): “A Theory of Conformity.” Journal of Political Economy, 102(5).

[10] Bernheim, Douglas and Antonio Rangel (2004). “Cue-Triggered Decision Processes.” American Economic Review, 94:5, p. 1558-1590.

[11] Bikhchandani, Sushil, David Hirshleifer and Ivo Welch (1992). “A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades.” Journal of Political Economy, 100(5).

[12] Bisin, Alberto, Andrea Moro and Giorgio Topa (2011). “The empirical content of models with multiple equilibria in economies with social interactions.” National Bureau of Economic Research, No. w17196.

[13] Bloomberg U.S. National Poll (2014). Retrieved from http://www.bloomberg.com/infographics/2014-03-12/national-poll.html#gay-rights.

[14] Bolton, Lisa and Joseph Alba (2012). “When Less is More: Consumer Aversion to Unused Utility.” Journal of Consumer Psychology, 22:3, p. 369-383.

[15] Brock, William A. and Steven N. Durlauf (2001): “Discrete Choice with Social Interactions.” Review of Economic Studies, 68(2).

[16] Brownell, Kelly and Kenneth Warner (2009). “The Perils of Ignoring History: Big Tobacco Played Dirty and Millions Died. How Similar is Big Food?” Milbank Quarterly, 87:1, p. 259-294.

[17] DellaVigna, Stefano, John A. List, Ulrike Malmendier and Gautam Rao (2013): “Voting to Tell Others.” Working Paper.

[18] Diekmann, Andreas, Wojtek Przepiorka and Heiko Rauhut (2011). "Lifting the Veil of Ignorance: An Experiment on the Contagiousness of Norm Violations." Nuffield Centre for Experimental Social Sciences Discussion Paper Series, 2011-004.

[19] Dixit, Avinash (2003). "Clubs with Entrapment." American Economic Review, 93(5).

[20] Dybvig, Philip H. and Chester S. Spatt (1983): "Adoption Externalities as Public Goods." Journal of Public Economics, 20(2), 231-247.

[21] Easley, David and Jon Kleinberg (2010): *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*, Cambridge University Press.

[22] Farrell, Joseph and Paul Klemperer (2007): "Coordination and Lock-in: Competition with Switching Costs and Network Effects." *Handbook of Industrial Organization*, Vol 3.

[23] Farrell, Joseph and Garth Saloner (1985): "Standardization, Compatibility, and Innovation." RAND Journal of Economics, 16(1), 70-83.

[24] Festinger, Leon (1962). *A Theory of Cognitive Dissonance*. Stanford, California: Stanford University Press.

[25] Fogli, Alessandra and Laura Veldkamp (2011). "Nature or Nurture? Learning and the Geography of Female Labor Force Participation." Econometrica, 79(4).

[26] Frankel, David M., Stephen Morris and Ady Pauzner (2003). "Equilibrium Selection in Global Game with Strategic Complementarities." Journal of Economic Theory, 108(1).

[27] Fudenberg, Drew and David K. Levine (2012). "Timing and Self-Control." Econometrica, 80:1, p. 1-42.

[28] Gallup (2007): "More Smokers Feeling Harassed by Smoking Bans." Retrieved from http://www.gallup.com/poll/28216/More-Smokers-Feeling-Harassed-Smoking-Bans.aspx

[29] Gallup (2013). Retrieved from http://www.gallup.com/file/poll/169868/REV_Views_on_Same_Sex_Marriage_140521.pdf

[30] Gallup (2014). Retrieved from http://www.gallup.com/poll/169640/sex-marriage-support-reaches-new-high.aspx

[31] Gerard, Harold, Wilhelmy, Roland, and Edward Conolley (1968): "Conformity and Group Size." Journal of Personality and Social Psychology, 8(1), 79-82.

[32] Granovetter, Mark (1978): "Threshold Models of Collective Behavior." American Journal of Sociology, 83(6).

[33] Guiso, Luigi, Sapienza, Paola, and Luigi Zingales (2009): "Moral and Social Constraints to Strategic Default on Mortgages." Working Paper.

[34] Gul, Faruk and Wolfgang Pesendorfer (2001). "Temptation and Self-Control." Econometrica, 69:6, p. 1403-1435.

[35] Gul, Faruk and Wolfgang Pesendorfer (2007). "Harmful Addiction." Review of Economic Studies, 74, p. 147-172.

[36] Ling, Pamela and Stanton Glantz (2002). "Why and How the Tobacco Industry Sells Cigarettes to Young Adults: Evidence from Industry Documents." American Journal of Public Health, 92:6, p. 908-916.

[37] Milgrom, Paul and John Roberts (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities." Econometrica, 58(6), 1255-1277.

[38] Minnesota Population Center and State Health Access Data Assistance Center, Integrated Health Interview Series: Version 4.0. Minneapolis: University of Minnesota, 2011.

[39] Noelle-Neumann, Elisabeth (1974): "The Spiral of Silence: A Theory of Public Opinion." Journal of Communication, 24(2).

[40] Pollak, Robert (1970). "Habit Formation and Dynamic Demand Functions." Journal of Political Economy, 78:4, p. 745-763.

[41] Rolls, Barbara, Erin Morris and Liane Roe (2002). "Portion size of food affects energy intake in normal-weight and overweight men and women." American Journal of Clinical Nutrition 76:6, p. 12071213.

[42] Saenz-de-Miera, Belen et al. (2010). "Self-reported Price of Cigarettes, Consumption, and Compensatory Behaviors in a Cohort of Mexican Smokers Before and After a Cigarette Tax Increase." Tobacco Control, vol. 19, p. 481-487.

[43] Schmeidler, David (1973). "Equilibrium Points of Nonatomic Games." Journal of Statistical Physics, vol. 7, 4:295-300.

[44] Scragg, Robert et al. (2008). "Diminished Autonomy over Tobacco can Appear with the First Cigarettes." Addictive Behaviors, doi:10.1016/j.addbeh.2007.12.002

[45] Schelling, Thomas C. (1973): "Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices with Externalities." The Journal of Conflict Resolution, 17(3).

[46] Schelling, Thomas C. (1978): *Micromotives and Macrobehavior*

[47] Shaw, Donald, Maxwell McCombs, David Weaver, and Bradley Hamm (1999): "Individuals, Groups, and Agenda Melding: A Theory of Social Dissonance." International Journal of Public Opinion Research, 11(1).

[48] Smith, Katherine Clegg et al. (2007). "Buying and Selling 'Loosies' in Baltimore: The Informal Exchange of Cigarettes in the Community Context." Journal of Urban Health: Bulletin of the New York Academy of Medicine, 84:4, p. 494-507.

[49] Thrasher, James F. et al. (2009). "Does the Availability of Single Cigarettes Promote of Inhibit Cigarette Consumption? Perceptions, Prevalence, and Correlates of Single Cigarette Use among Adult Mexican Smokers." Tobacco Control, vol. 18, p. 431-437.

[50] Thrasher, James F. et al. (2011). "Consumption of Single Cigarettes and Quitting Behavior: A Longitudinal Analysis of Mexican Smokers." BioMed Central Public Health, 11:134.

[51] Topkis, Donald M. (1979). "Equilibrium Points in Nonzero-Sum n-Person Submodular Games." Siam Journal on Control and Optimization, 17.6: 773-787.

[52] Vives, Xavier (1990). "Nash Equilibrium with Strategic Complementarities." Journal of Economic Theory, vol. 134, 339-360.

[53] Vlosky, Denese Ashbaugh and Pamela A. Monroe (2002). "The Effective Dates of No-Fault Divorce Laws in the 50 States." Family Relations 51.4: 317-324.

[54] Wansink, Brian, James Painter and Jill North (2005). "Bottomless Bowls: Why Visual Cues of Portion Size May Influence Intake." Obesity Research, 13:1, p. 93-100.

[55] Yang, Jian and Xiangtong Qi (2013) "The Nonatomic Supermodular Game." Games and Economic Behavior, vol. 82, 609-620.

Appendix A

Chapter 2

A.1 Proofs

A.1.1 General Two-Action Games

Lemma 24. *Suppose s is an equilibrium, $\Delta u^i = \Delta u^j$ and $s^i \neq s^j$. For any two agents with $\Delta u^k = \Delta u^{k'} \neq \Delta u^i$, $s^k = s^{k'}$.*

Proof. Suppose without loss of generality that $s^i = A$ and $s^j = B$. If $\Delta u^k > \Delta u^j$ then $u^k(B) - u^k(A) = u^{k'}(B) - u^{k'}(A) > u^j(B) - u^j(A)$ so by monotonicity, $s^k = B$ and $s^{k'} = B$. Likewise, if $\Delta u^k < \Delta u^j$ then $u^k(A) - u^k(B) = u^{k'}(A) - u^{k'}(B) > u^j(A) - u^j(B)$ and it must be that $s^k = s^{k'} = A$. □

A.1.2 Two-Type Games

The following proofs are for the game with agents of only two different values of Δu^i. Assume throughout that $u^a(A) > u^a(B)$ and $u^b(B) > u^b(A)$.

Existence of Equilibria in Two-Type Game:

The following discussion gives the set of equilibria for different configurations of the game and also provides a sketch of an existence proof done by enumeration of all possibilities. There are three [1] possible configurations of $\{\Delta u^a, \Delta u^b, d\}$, as shown in the table below.

If the social dissonance of being the only person to choose that action is outweighed by the gap in intrinsic utility between an agent's preferred and non-preferred actions, the agent will always choose true to type; it is a dominant strategy. If this is so for both types of agents as in case (1), then only a true-to-type equilibrium is possible, since the best response for all agents is to choose their intrinsically-preferred action regardless of the actions of others.

In case (2), there is one type that has a dominant strategy of choosing true to its intrinsic preference. If there are not enough agents of the type with the weaker preference, then the true-to-type conditions do not hold and the only equilibrium is monolithic in the action of the dominant type.

In case (3), both monolithic-A and monolithic-B equilibria always exist. In addition, there is a true-to-type equilibrium if the true-to-type conditions hold. The table summarizes the existence of equilibria in each case. "ttt" denotes existence if and only if the true-to-type conditions hold.

Case	Parameters	True-to-Type	Monolithic-A	Monolithic-B
1	$\Delta u^a > d(A,0)$ and $\Delta u^b > d(B,0)$	Yes	No	No
2A	$\Delta u^a > d(A,0)$ and $\Delta u^b < d(B,0)$	ttt	Yes	No
2B	$\Delta u^a < d(A,0)$ and $\Delta u^b > d(B,0)$	ttt	No	Yes
3	$\Delta u^a < d(A,0)$ and $\Delta u^b < d(B,0)$	ttt	Yes	Yes

Monotonicity

Claim For the two-action, two-type game, in equilibrium, at least one agent of one type plays against-type, then all agents of the other type play true-to-type.

[1] There are more cases where Δu^a or Δu^b might be exactly equal to $d(A,0)$ or $d(B,0)$ respectively. These are left to the reader.

Proof. Suppose s is an equilibrium where $\hat{N}_b(s)$ type-b agents choose B, with $\hat{N}_b(s) \in \{0, 1, 2, ..., N_b\}$ and at least one type-a plays action B, so $\hat{N}_a(s) > 0$ and $\hat{N}_a(s) < N_a$. $\exists i \in \{1, 2, ..., N_a\}$ who chooses B and therefore has $u^a(A) - d(A, N_{s^{-i}}(A)) \leq u^a(B) - d(B, N_{s^{-i}}(B))$. Since $u^a(A) > u^a(B)$, we must have $d(A, N_{s^{-i}}(A)) > d(B, N_{s^{-i}}(B))$ for this agent i.

Now, from the perspective of any other agent j, the tally of other agents who choose B, $N_{-j}(B)$ may exclude agent j but it will definitely include agent i, whereas agent i was not included in the tally $N_{-i}(B)$. That is, $N_{-j}(B) \geq N_{-i}(B)$. Likewise, $N_{-j}(A) \leq N_{-i}(A)$.

This makes the decision of a type-b agent easy. For any agent $j \neq i$, $d(A, N_{s^{-j}}(A)) \geq d(A, N_{s^{-i}}(A))$ and $d(B, N_{s^{-j}}(B)) \leq d(B, N_{s^{-i}}(B))$ so $d(A, N_{s^{-j}}(A)) > d(B, N_{s^{-j}}(B))$. For a type-b, $u^b(B) > u^b(A)$ so we will have

$$u^b(B) - d(B, N_{s^{-j}}(B) > u^b(A) - d(A, N_{s^{-j}}(A))$$

That is, when in equilibrium type-a's are mixing or all playing B, it follows that a type-b's only best response is to choose B. We cannot have mixing for both type-a and type-b. Similar logic applies for the case where type-b mixes; if that is so, then we must have $\hat{N}_a(s) = N_a$. □

A.2 Figures

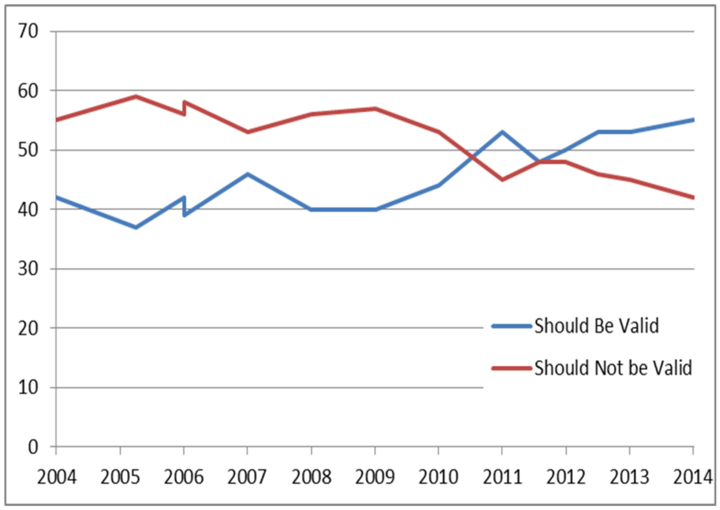

Figure A.1: Survey Responses to "Do you think marriages between same-sex couples should or should not be recognized by the law as valid, with the same rights as traditional marriages?" (source: Gallup)

Appendix B

Chapter 3

B.1 Proofs

Corollary 3: In the two-action, two-type game, if in equilibrium some agents of one type play against type, then all agents of the other type must play true-to-type.

Proof. Assume that $u^a(A) > u^a(B)$ and $u^b(B) > u^b(A)$. Suppose in equilibrium, under action profile s, that some type-a agents choose A and some type-b agents choose B. Denote by $\hat{\mu_a}(s)$ the measure of type-a agents who choose A, and by $\hat{\mu_b}(s)$ the measure of type-b agents who choose B. Suppose further that some type-a's choose B in this equilibrium. The measure of those agents is $\mu_a - \hat{\mu_a}(s) > 0$. The measure of type-b's that choose A is $\mu_b - \hat{\mu_b}(s)$, though we have not determined whether this is positive or not.

If some type-a choose B then

$$u^a(B) - d(B, \mu_a - \hat{\mu_a}(s) + \hat{\mu_b}(s)) \geq u^a(A) - d(A, \hat{\mu_a}(s) + (\mu_b - \hat{\mu_b}(s)))$$

Since $u^a(A) > u^a(B)$, in order for the inequality above to hold, it must be that

$$d(A, \hat{\mu_a}(s) + (\mu_b - \hat{\mu_b}(s))) > d(B, \mu_a - \hat{\mu_a}(s) + \hat{\mu_b}(s)) \quad \text{(B.1)}$$

This, combined with the fact that $u^b(B) > u^b(A)$ implies that

$$u^b(B) - d(B, \mu_a - \hat{\mu_a}(s) + \hat{\mu_b}(s)) \geq u^b(A) - d(A, \hat{\mu_a}(s) + (\mu_b - \hat{\mu_b}(s)))$$

Hence, only B is a best response for type-b agents. An analogous argument can be made to show that if some type-b's choose A in equilibrium, then only A is a best response for type-a agents. □

Appendix C

Chapter 4

C.1 Additional Proofs

Lemma 25. *Consumers facing a weak impulse can have preference ranking $S > L > \emptyset$.*

Proof. It must be that $-U(S) < -U(L) < -U(\emptyset)$. In order to have this, it must be that:

$$p_L + H + \mathbb{P}_R(T)H + \mathbb{P}_R(W)H < I_W + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W \tag{C.1}$$

and

$$p_S + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W < p_L + \mathbb{P}_R(T)H + \mathbb{P}_R(W)H \tag{C.2}$$

Let $G = \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W - (\mathbb{P}_R(T)H + \mathbb{P}_R(W)H)$. By Assumption 13, $p_L + H > I_W$, so it must be that $G > p_L + H - I_W$ in order for (C.1) to hold. Meanwhile, $p_L > p_S$ so it must be that $G < p_L - p_S$. for (C.2) to hold. By Assumption 13, $H - I_W < -p_S$, so G is contained in a non-empty interval $(p_L + H - I_W, p_L - p_S)$ and there are not any restrictions on the relationship between $\mathbb{P}_\emptyset$ and $\mathbb{P}_R$ that preclude G from being in this interval.

□

Lemma 26. *Consumers facing a weak impulse can have preference ranking $S > \emptyset > L$.*

Proof. It must be that $-U(S) < -U(\emptyset) < -U(L)$. In order to have this, it must be that:

$$p_S + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W < I_W + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W \tag{C.3}$$

and

$$I_W + \mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W < p_L + H + \mathbb{P}_R(T)H + \mathbb{P}_R(W)H \tag{C.4}$$

$I_W > p_S$ by assumption 13 so (C.3) is always true. Also by Assumption 13, $p_L + H - I_W > 0$ so we need only $\mathbb{P}_\emptyset(T)I_T + \mathbb{P}_\emptyset(W)I_W - (\mathbb{P}_R(T)H + \mathbb{P}_R(W)H) < p_L + H - I_W$ for (C.4) to be true. There are no restrictions on the relationship between $\mathbb{P}_\emptyset$ and $\mathbb{P}_R$ that preclude this. This is simply a bound on how costly expected unfulfilled temptation-phase impulses are relative to the expected health cost of fulfilled temptations. □

C.2 Expected Payoffs

Here, I write out in longhand the expected costs associated with each menu choice when facing different impulses.

C.2.1 One Period Model

When facing I_T, the total costs of purchasing each menu and consuming optimally are:
$\emptyset$: $I_T + P_1(T)I_T + P_1(W)I_W$
S: $p_S + H + P_1(T)I_T + P_1(W)I_W$
L: $p_L + H[1 + P_2(T) + P_2(W)]$

When facing I_W, the total costs of purchasing each menu and consuming optimally are:
$\emptyset$: $I_W + P_1(T)I_T + P_1(W)I_W$

S: $p_S + H + P_1(T)I_T + P_1(W)I_W$

L: $p_L + H[1 + P_2(T) + P_2(W)]$

C.2.2 Two Period Model

When $\emptyset, S, L$ are all available,

$V(k) = k(\mathbb{P}(T_1) + \mathbb{P}(W_1))min\{p_L + H + a_R[\mathbb{P}(T_2)H + \mathbb{P}(W_2)H], p_S + H + \emptyset[\mathbb{P}(T_2)I_T + \mathbb{P}(W_2)I_W]\}$

$V(K) = K(\mathbb{P}(T_1) + \mathbb{P}(W_1))min\{p_L + H + a_R[\mathbb{P}(T_2)H + \mathbb{P}(W_2)H], p_S + H + \emptyset[\mathbb{P}(T_2)I_T + \mathbb{P}(W_2)I_W]\}$

$V(\hat{K}) = \hat{K}(\mathbb{P}(T_1) + \mathbb{P}(W_1))min\{p_L + H + a_R[\mathbb{P}(T_2)H + \mathbb{P}(W_2)H], p_S + H + \emptyset[\mathbb{P}(T_2)I_T + \mathbb{P}(W_2)I_W]\}$

When only $\emptyset, L$ are available,

$V(k) = k\mathbb{P}(T_1)min\{p_L + H + a_R[\mathbb{P}(T_2)H + \mathbb{P}(W_2)H], I_T + \emptyset[\mathbb{P}(T_2)I_T + \mathbb{P}(W_2)I_W]\}$
$+ k\mathbb{P}(W_1)min\{p_L + H + a_R[\mathbb{P}(T_2)H + \mathbb{P}(W_2)H], I_W + \emptyset[\mathbb{P}(T_2)I_T + \mathbb{P}(W_2)I_W]\}$

$V(K) = K\mathbb{P}(T_1)min\{p_L + H + a_R[\mathbb{P}(T_2)H + \mathbb{P}(W_2)H], I_T + \emptyset[\mathbb{P}(T_2)I_T + \mathbb{P}(W_2)I_W]\}$
$+ K\mathbb{P}(W_1)min\{p_L + H + a_R[\mathbb{P}(T_2)H + \mathbb{P}(W_2)H], I_W + \emptyset[\mathbb{P}(T_2)I_T + \mathbb{P}(W_2)I_W]\}$

$V(\hat{K}) = \hat{K}\mathbb{P}(T_1)min\{p_L + H + a_R[\mathbb{P}(T_2)H + \mathbb{P}(W_2)H], I_T + \emptyset[\mathbb{P}(T_2)I_T + \mathbb{P}(W_2)I_W]\}$
$+ \hat{K}\mathbb{P}(W_1)min\{p_L + H + a_R[\mathbb{P}(T_2)H + \mathbb{P}(W_2)H], I_W + \emptyset[\mathbb{P}(T_2)I_T + \mathbb{P}(W_2)I_W]\}$

When there is no consumption in period 1, probabilities of intense and weak impulses are $k\mathbb{P}(T_1)$ and $k\mathbb{P}(W_1)$. When either of these impulses is realized, the agent uses the decision rules in lemmas 21 and 22. If a negligible impulse is realized, $\emptyset$ is chosen.

When facing intense impulse I_T in period 1, the expected costs of each possible menu and consumption plan are:

$$
\begin{array}{ll}
\emptyset: & I_T + \emptyset[\mathbb{P}(T_1)I_T + \mathbb{P}(W_1)I_W] + \beta V(k) \\
S: & p_S + H + \emptyset[\mathbb{P}(T_1)I_T + \mathbb{P}(W_1)I_W] + \frac{K}{k}\beta V(k) \\
L\text{, resist } I_T \text{ and } I_W: & p_L + H + a_R[\mathbb{P}(T_1)I_T + \mathbb{P}(W_1)I_W] + \frac{K}{k}\beta V(k) \\
L\text{, resist } I_W: & p_L + H + a_R[\mathbb{P}(T_1)(H + \beta\frac{\hat{K}}{k}V(\hat{K})) + \mathbb{P}(W_1)(I_W + \beta\frac{K}{k}V(k))] \\
L\text{, give in:} & p_L + H + a_R[\mathbb{P}(T_1)H + \mathbb{P}(W_1)H] + \beta\frac{\hat{K}}{k}V(k)
\end{array}
$$

When facing weak impulse I_W in period 1, the expected costs of each possible menu and consumption plan are:

$$
\begin{array}{ll}
\emptyset: & I_W + \emptyset[\mathbb{P}(T_1)I_T + \mathbb{P}(W_1)I_W] + \beta V(k) \\
S: & p_S + H + \emptyset[\mathbb{P}(T_1)I_T + \mathbb{P}(W_1)I_W] + \beta\frac{K}{k}V(k) \\
L\text{, resist } I_T \text{ and } I_W: & p_L + H + a_R[\mathbb{P}(T_1)I_T + \mathbb{P}(W_1)I_W] + \beta\frac{K}{k}V(k) \\
L\text{, resist } I_W: & p_L + H + a_R[\mathbb{P}(T_1)(H + \beta\frac{\hat{K}}{k}V(k)) + \mathbb{P}(W_1)(I_W + \beta\frac{K}{k}V(k))] \\
L\text{, give in:} & p_L + H + a_R[\mathbb{P}(T_1)H + \mathbb{P}(W_1)H] + \beta\frac{\hat{K}}{k}V(k)
\end{array}
$$